

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2003, 06:36:14 ; Search time 2944 Seconds

(without alignments)
10369.852 Million cell updates/sec

Title: US-09-896-186B-23

Perfect score: 1049

Sequence: 1 accaagcatatatttattt.....tttaacgctccagaactag 1049

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:*

1: gb_ba:*

2: gb_htg:*

3: gb_in:*

4: gb_com:*

5: gb_ov:*

6: gb_ph:*

7: gb_pl:*

8: gb_pr:*

9: gb_ro:*

10: gb_sy:*

11: gb_un:*

12: gb_vl:*

13: gb_vl:*

14: gb_vl:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_cm:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_pl:*

25: em_ro:*

26: em_sts:*

27: em_un:*

28: em_vl:*

29: em_vl:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_pln:*

35: em_htg_rtd:*

36: em_htg_mam:*

37: em_htg_vrt:*

38: em_sy:*

39: em_htgo_hum:*

40: em_htgo_mus:*

41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1049	100.0	1049	6	AX375195	AX375195 Sequence
2	1015	96.8	1137	6	ATB40476	AJ404476 Arabidops
3	831	79.2	942	6	AX375173	AX375173 Sequence
4	302.2	28.8	118718	8	ATF18A5	AL035528 Arabidops
5	302.2	28.8	199667	8	ATCHRIV37	AL151537 Arabidops
6	151.2	14.4	130299	8	AC019012	AC019012 Genomic S
7	142	13.5	82484	8	AC002341	AC002341 Arabidops
8	111	10.6	4299	6	AX375189	AX375189 Sequence
9	111	10.6	5189	6	AX333649	AX333649 Sequence
10	111	10.6	5189	6	E16596	E16596 Human WS mr
11	111	10.6	5189	9	HUMDR	L76937 Homo sapien
12	111	10.6	5208	9	AF091214	AF091214 Homo sapi
13	110.6	10.5	5143	5	AF067418	AF067418 Xenopus l
14	97.8	9.3	200576	8	ATFOAO	Z97335 Arabidops
15	82.8	7.9	6476	6	AR149285	AR149285 Sequence
16	82.8	7.9	6476	10	AF241636	AF241636 Mus muscu
17	81.8	7.8	4201	10	MM097045	U97045 Mus musculu
18	81.2	7.7	4955	10	D86527	D86527 Mouse werne
19	81.2	7.7	5034	10	D86526	D86526 Mouse werne
20	81.2	7.7	5050	10	AF091215	AF091215 Mus muscu
21	81.2	7.7	5058	6	E16141	E16141 CDNA encodl
22	65.4	6.2	155597	8	OSJN00161	AL662964 Oryza sat
23	55.4	5.3	473	11	G73264	G73264 csnprn-pcr
24	55.4	5.3	98844	9	HSWRNG1	AF181896 Homo sapl
25	55.4	5.3	180098	9	AC084736	AC084736 Homo sapl
26	51.2	4.9	97348	10	AF091216	AF091216 Mus muscu
27	51.2	4.9	227725	2	AC115809	AC115809 Mus muscu
28	49.2	4.7	7218	6	I66494	I66494 Sequence 14
29	48.2	4.6	174993	2	AC098510	AC098510 Rattus no
30	46	4.5	168698	2	OSJN00023	AL606588 Oryza sat
31	46	4.4	2865	10	BC018508	BC018508 Mus muscu
32	45.2	4.3	434	11	G73268	G73268 csnprn-pcr
33	45	4.3	236120	14	AF063866	AF063866 Melanopl
34	44.4	4.2	2077	9	BC001962	BC001962 Homo sapl
35	44.4	4.2	2833	9	AK001600	AK001600 Homo sapl
36	44.4	4.2	2844	9	HSMB05527	AL834434 Homo sapl
37	44.4	4.2	3404	9	AK056025	AK056025 Homo sapl
38	44.4	4.2	198421	2	AC125180	AC125180 Mus muscu
39	44.2	4.2	164679	2	AC096680	AC096680 Pan trogl
40	43.4	4.1	125020	9	AF429315	AF429315 Homo sapl
41	43	4.1	115489	2	AC117072	AC117072 Dictyoste
42	42.8	4.1	408	6	AX300942	AX300942 Sequence
43	42.6	4.1	116660	9	AC002074	AC002074 Homo sapl
44	42.4	4.0	160780	2	AL445653	AL445653 Homo sapl
45	42.2	4.0	1141	6	AX083744	AX083744 Sequence

ALIGNMENTS

RESULT 1

AX375195

LOCUS AX375195 1049 bp DNA linear PAT 01-MAR-2002

DEFINITION Sequence 23 from Patent WO0210562.

ACCESSION AX375195

VERSION AX375195.1 GI:19169935

KEYWORDS

SOURCE

ORGANISM

thale cress.

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE

1 Levin,J.Z., Phillips,K.L., Budziszewski,G.J., Meins,F.J. and Glazov,E.A.

TITLE Methods of controlling gene expression and gene silencing
JOURNAL Patent: WO 0210362-A 23 07-FEB-2002;
Syngenta Participations AG (CH) ; Novartis Forschungsf Stiftung
Zweigniederlassung (CH)
FEATURES
source
1.1049
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
BASE COUNT 292 a 216 c 220 g 321 t
ORIGIN

Query Match 100.0%; Score 1049; DB 6; Length 1049;
Best Local Similarity 100.0%; Pred. No. 1.3e-259;
Matches 1049; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCAAGCATTAATTTTATTTTGTTCAGTAAAGAAATGTCATCGTCAATTTGA 60
DB 1 ACCAAGCATTAATTTTATTTTGTTCAGTAAAGAAATGTCATCGTCAATTTGA 60
QY 61 TCGACGACGCTTTACAGAGAGAGCTTCGCTATCGAGCCGATCGAAGCTTCTACA 120
DB 61 TCGACGACGCTTTACAGAGAGAGCTTCGCTATCGAGCCGATCGAAGCTTCTACA 120
QY 121 ATTTCGCCGCTTCTTCTTCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
DB 121 ATTTCGCCGCTTCTTCTTCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
QY 181 CCGTCATGCGCAGAGAGAGATCCAAATCAATCCCAATTAATTCGCGCAATTCG 240
DB 181 CCGTCATGCGCAGAGAGAGATCCAAATCAATCCCAATTAATTCGCGCAATTCG 240
QY 241 CTGCTTCATCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 300
DB 241 CTGCTTCATCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 300
QY 301 GGAATTTTCCAGAAAGAGGTGGTGGTGAATTTTGTATACGAAGCTGCTACAGG 360
DB 301 GGAATTTTCCAGAAAGAGGTGGTGGTGAATTTTGTATACGAAGCTGCTACAGG 360
QY 361 TTGATTAAGCAGCAATGACACTTATTAAGTCTTGTATACCAAGAGATGAATCTGAA 420
DB 361 TTGATTAAGCAGCAATGACACTTATTAAGTCTTGTATACCAAGAGATGAATCTGAA 420
QY 421 TACCTTTTGGCTTGGATTTGATGAGGAGACCAATTTTAAAGAGTGTCTCCGG 480
DB 421 TACCTTTTGGCTTGGATTTGATGAGGAGACCAATTTTAAAGAGTGTCTCCGG 480
QY 481 GGAAGGTGGACGTGACATATGATATGATATTTTGTGATGATGATGATTTT 540
DB 481 GGAAGGTGGACGTGACATATGATATGATATTTTGTGATGATGATGATTTT 540
QY 541 TTGATTTGCTATCCCTCAAGTCTCCACATCTTATGAAGATTCMAACTGTAAAG 600
DB 541 TTGATTTGCTATCCCTCAAGTCTCCACATCTTATGAAGATTCMAACTGTAAAG 600
QY 601 TAGGTATGGAATGATGAGTCTGTGAAGCTTTTCATGACTATGAGAGTTAGTACA 660
DB 601 TAGGTATGGAATGATGAGTCTGTGAAGCTTTTCATGACTATGAGAGTTAGTACA 660
QY 661 AAGATGTGAGATCTTTCAGATTAGCCAAACCAAAATTTGGTGAGATTAATAATGG 720
DB 661 AAGATGTGAGATCTTTCAGATTAGCCAAACCAAAATTTGGTGAGATTAATAATGG 720
QY 721 GCCTTCCTCACTACTGAGACACTTGTGCAAGAGCTCTGAAGCCCAACAGATTA 780
DB 721 GCCTTCCTCACTACTGAGACACTTGTGCAAGAGCTCTGAAGCCCAACAGATTA 780
QY 781 GCGTTGGGAAGTGGAGTTTATCTCTGTCAAGAGAGATTAACATTAACGACGAG 840
DB 781 GCGTTGGGAAGTGGAGTTTATCTCTGTCAAGAGAGATTAACATTAACGACGAG 840
QY 841 ATGCTTATGCTTATGAGCATCTTACAGAGTTTAAAGACCTTCTGATGCTGACGTG 900
DB 841 ATGCTTATGCTTATGAGCATCTTACAGAGTTTAAAGACCTTCTGATGCTGACGTG 900

DB 841 ATGCTTATGCTTATGAGCATCTTACAGAGTTTAAAGACCTTCTGATGCTGACGTG 900
QY 901 GCTCATTAAGTGAAGAGAGAGCTTAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 960
DB 901 GCTCATTAAGTGAAGAGAGAGCTTAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 960
QY 961 ATGATATGATACACCTTATCTAGTCAAGTATGATGATGATGATGATGATGATGATG 1020
DB 961 ATGATATGATACACCTTATCTAGTCAAGTATGATGATGATGATGATGATGATGATG 1020

QY 1021 TCTGTCCTTTTAAAGTCCGCAAGACTAG 1049
DB 1021 TCTGTCCTTTTAAAGTCCGCAAGACTAG 1049

RESULT 2
ATH404476 1137 bp mRNA linear PLN 02-NOV-2000
LOCUS
DEFINITION Arabidopsis thaliana mRNA for exonuclease (wrxexo gene).
ACCESSION AJ404476
VERSION AJ404476.1 GI:1121454
KEYWORDS exonuclease; wrxexo gene.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; Rosidae; eustosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
AUTHORS Hartung, F., Plachova, H. and Puchta, H.
TITLE Molecular characterisation of RecQ homologues in Arabidopsis thaliana
JOURNAL Nucleic Acids Res. 21, 4275-4282 (2000)
REFERENCE
AUTHORS Hartung, F., Plachova, H. and Puchta, H.
TITLE Molecular characterisation of RecQ homologues in Arabidopsis thaliana
JOURNAL Unpublished
REFERENCE
AUTHORS Hartung, F.
TITLE Direct Submission
JOURNAL Submitted (13-JUN-2000) Hartung F., Cyto genetics, Institute of plant genetics and crop plant research (IPK), Corrensstr. 3, Sechen Anhalt, 06466 Gatersleben, GERMANY
REMARK
COMMENT Revised by author 14-JUN-2000
FEATURES
location/Qualifiers
1.1137
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/chromosome="4"
/issue_type="flowers"
36..893
/gene="wrxexo"
36..893
/gene="wrxexo"
/function="putative role in replication and recombination"
/codon_start=1
/product="exonuclease"
/protein_id="CA14871.1"
/db_xref="GI:1121454"
/translation="MSSSWIDAFTFEELIAIDAIEASYNFSRSSSSSSAAPTVOA
TTSVHCHIEDPNOIPNNIRDLPRSTSTSYRPLSRANRPFAMRGRILYSK
TAVEVDRKAMOLIKVLDTRDESGIAFVGLDWRSPFRGVLPKVAATVQICVDSNY
CDVMHIFHSGIPOSLOHLIEDSTLVVVGISIDGSDVKLPHDVGVSITKVEDLSDLNO
KIGDKKMGIASLITELVCKELKPRNIRIRLNNEFTPLSQOLQYATDAVYASWHLK
DLPDAVSGS"
BASE COUNT 318 a 226 c 238 g 355 t
ORIGIN

Query Match 96.8%; Score 1015; DB 8; Length 1137;
Best Local Similarity 99.1%; Pred. No. 7.6e-251;
Matches 1034; Conservative 0; Mismatches 0; Indels 9; Gaps 1;

Db 601 GACTATGAGCTTGTGATCAAGATGTTGAGGATCTTTCAGATTTAGCCAAACCAAAATT 660
 QY 702 GGTGAGATTAATAATGCGGCTTGGCTCACTAAGTACAGACACTTGTTCAGAAAGACTC 761
 Db 661 GGTGAGATTAATAATGCGGCTTGGCTCACTAAGTACAGACACTTGTTCAGAAAGACTC 720
 QY 762 CTGAGCAACAGATCAGCTGGAGCTGGAGCTTTATCTCTGTCAAGCAGAC 821
 Db 721 CTGAGCAACAGATCAGCTGGAGCTGGAGCTTTATCTCTGTCAAGCAGAC 780
 QY 822 TTTCATATGACGACAGGATGCTTATGCTTCAATGAGCTTTACAGAGTTCTTAAGAC 881
 Db 781 TTTCATATGACGACAGGATGCTTATGCTTCAATGAGCTTTACAGAGTTCTTAAGAC 840
 QY 882 CTTCGATGCTGTCACTGCTCATTAACGTAAGAGAGACCTTAA 928
 Db 841 AAACCATCTCTCACACTCAACGACCTTGAAGCAAAATTCACACA 887

 RESULT 4
 ATFL8A5/c
 LOCUS
 DEFINITION 118718 bp DNA linear PLN 18-AUG-1999
 Arabidopsis thaliana DNA chromosome 4, BAC clone F18A5 (ESSA project).
 ACCESSION AL035528
 VERSION AL035528.2 GI:5748493
 KEYWORDS
 SOURCE
 ORGANISM
 Arabidopsis thaliana.
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 1 (bases 1 to 118718)
 Bevan, M., Weber, N., Grueninger, D., Schmidheini, T., Bancroft, I.,
 Mewes, H.W., Mayer, K.F.X., Lemcke, K. and Schueller, C.
 Unpublished
 2 (bases 1 to 118718)
 EU Arabidopsis sequencing project.
 Direct Submission
 Submitted (18-AUG-1999) MIPS, at the Max-Planck-Institut fuer
 Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
 schueller@mpi-biochem.mpg.de, mayer@mpi-biochem.mpg.de, Project
 Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
 Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
 E-mail: michael.bevan@bbsrc.ac.uk
 On Aug 20, 1999 this sequence version replaced gi:4455290.
 Information on performance of analysis and a more detailed
 annotation of this entry and other sequences of chromosomes 3, 4
 and 5 can be viewed at: <http://www.mips.biochem.mpg.de/proj/thal/>.

 COMMENT
 FEATURES
 source
 1. 118718
 /organism="Arabidopsis thaliana"
 /variety="Columbia"
 /db_xref="taxon:3702"
 /chromosome="4"
 misc-feature
 1. 3993
 /note="basepairs 1-3993 are not integral part of clone
 F18A5 but are part of BAC FT23H4"
 3154..4320
 /gene="F18A5.10"
 3154..4320
 /gene="F18A5.10"
 /note="weak similarity to CADMIUM-INDUCED PROTEIN AS30 -
 Arabidopsis thaliana, GB:P42736"
 Arabidopsis thaliana, GB:P42736"
 /codon_start=1
 /product="putative protein"
 /protein_id="CA36826.2"
 /db_xref="GI:5748494"
 /translation="MTPPIHTQSHLIVYINISPYLSKLTGFIILMTWTQTKNKR
 NMEDPKIETSMHDLSSGIYGLSSIPOLGVPILFEGMKSPILPASSTSY
 FVSPHDLHSSIHSPVAVSPNPLFESSPOSDPHHPSKPNILFLFKEPKLLLS
 QSSNMSPYHKYIPNSFYOSDONRNEMVEINKLITLTPSPSGRNYLSTTKTPAKSK
 TRKVVOTTPTKLIRGVROHNMKVAEIRLPNRRVWLGTFETAEQAAAVDTAAV

ILRGEFAHLNFPDLKLQKSGSLRCMIASILESKIQIQLSSQVSNPSPPPVGVGPE
 QKNHMKMESGEVVMKKOKSRKREVEMGDEVQLSRMPSLMDLIMALSPHSS"
 3154..4320
 /gene="F18A5.10"
 /number=1
 5109..7045
 /gene="F18A5.20"
 complement(join(5109..5150,5302..6455,6496..7045))
 /gene="F18A5.20"
 complement(join(5109..5150,5302..6455,6496..7045))
 /gene="F18A5.20"
 /note="similarity to IFA-binding protein, Arabidopsis
 thaliana, AF004356"
 /codon_start=1
 /product="putative protein"
 /protein_id="CA36827.1"
 /db_xref="GI:4455292"
 /translation="MRQEVKSMTFKCLVNAFIDLSVAEELCASFIYVYTSKFLIF
 GLDIPCPDGLISECFQESLRLNLPVKIISVQSKNRTFFDSILINGKRCERRR
 VOLEDEVSTTPSVGKFNKASGFDLTAGSLKRGKSVKSLSPFRSPYGFKNHQ
 SCGLKSFEGSYDENPDLVNSVNVDFDASCNIVSGKALEVSLKSVLSVGGC
 VGGACSSPGMVORTVYMSQVIGEEAASALALELERRNAASAADALGILRL
 QERKASTEMARQYORMIERKSAFDAEKSILKEILRRERHFLKEVDYROMFL
 ETQPHNTPDSKPAQIERLOTPOQITPPMDMETVYVSSGEIETFNQMDTEQDSRS
 EPPDAUSDILKREBGETLYAEIVSTSDIAVSKLCEPHIDCHVIDHIVYTDSD
 NKQLNVPSDHATQDLKIDRSQVSDTSYAFPGKSMSTNMNRNSKALDYERLKE
 SEVGLNGRLRAVQKGREKISFSSKDSQSKOVORCGDKTSRFWEARRSGHLDSSPSS
 TVYKAMSPDLUSA"
 complement(5109..5150)
 /gene="F18A5.20"
 /number=1
 complement(5151..5301)
 /gene="F18A5.20"
 /number=1
 complement(5302..6455)
 /gene="F18A5.20"
 /number=2
 complement(6456..6495)
 /gene="F18A5.20"
 /number=2
 complement(6496..7045)
 /gene="F18A5.20"
 /number=3
 7880..9513
 /gene="F18A5.30"
 complement(join(7880..8293,8499..8568,8647..8725,
 8918..9055,9232..9255,9330..9513))
 /gene="F18A5.30"
 complement(join(7880..8293,8499..8568,8647..8725,
 8918..9055,9232..9255,9330..9513))
 /gene="F18A5.30"
 /note="contains EST gb:H76035, AA651479"
 /codon_start=1
 /product="putative protein"
 /protein_id="CA36828.1"
 /db_xref="GI:4455293"
 /translation="MYSAIRSRLPLDGLSDYDGTNLPIDACTLVLTTPKPLRWTS
 ELHERFDVATGCGPDISRKNLTATPTKPTIARTMGVKGTLTHLSKHQVLMPLPS
 ISNVASIALRNKQISCVAESODGSSSTSLRLAAGONSQVTELRNVOVORL
 HEOLVQRKQLRIEAGKYLQSLILEKACATIEQAVNAPGLEAREELSELAKRSI
 TNGCGGTSTFTDKMTPSLSELAVAIERKNCSSSSLSTSVGSPPVSAALMKRQ
 RGVFGNGDSVVVGHDAGVWPPSSISG"
 complement(7880..8293)
 /gene="F18A5.30"
 /number=1
 complement(8294..8498)
 /gene="F18A5.30"
 /number=1
 complement(8499..8568)
 /gene="F18A5.30"
 /number=2
 complement(8569..8646)
 /gene="F18A5.30"
 intron
 intron
 exon
 intron


```
/number=2
complement(8647. .8725)
/gene="F18A5.30"
/number=3
complement(8726. .8917)
/gene="F18A5.30"
/number=3
complement(8918. .9055)
/gene="F18A5.30"
/number=4
complement(9056. .9231)
/gene="F18A5.30"
/number=4
complement(9232. .9255)
/gene="F18A5.30"
/number=5
complement(9256. .9329)
/gene="F18A5.30"
/number=5
complement(9330. .9513)
/gene="F18A5.30"
/number=6
complement(9513. .9513)
/gene="F18A5.40"
complement(join(10627. .13689,13812. .13823))
/gene="F18A5.40"
complement(join(10627. .13689,13812. .13823))
/gene="F18A5.40"
/note="similarity to various predicted proteins,
Arabidopsis thaliana
contains cytochrome c family heme-binding site signature
[CNDOHA]"
/codon_start=1
/product="putative protein"
/protein_id="CAB36829.1"
/db_xref="GI:4455294"
/translation="MVFLLCGTRRASPAAISVYISEDESPQERKIDSENGKIRPNH
OTLKLKGLCTKNGSDEGRKHSOILKIGDSGCLSEKLFDFYLRKQDYGAFK
FDEMPERTFTNMKMKIKELASRLNGEYPLGFLVRYNSVNPBETFGVLEACGGS
VAFDVEQIHARILYQGLRSTVVCNPLIDLISRNGEYDLARVVDGLRLKHSSWA
MISGLSKNECEAIRLFCDMYVIGIMPTPYAFSSVLSACKKIELEIGEDJHGLV
LGSFSDYVCNALVSLYPLHGNLSAEHIFSMQRQDVTYNTLLNGSOGCGKAM
ELFRMHLDGLEPDSNTLASIVACSDGTLFRGOQLHAFTKLGFAFNNKLEGALN
LYACADIELTADLEFETEVENVVMYMLVAYGLDDLRSFRIFROMOLEIYPNO
YTSILKTKIRGLDELGEQIHSGIITNNQANVYCVSLIDMAKLGKIDTADIL
IRRAKDVSVTTMAGTYOTNFDDKALFTFRQMLDRGIRSDVEGLTNAVSCAGLQA
LKEGQIHQAQCVSGFSSDLFPQNALVTLYSRCKIEBSYLAFEQTEAGDNTAMALV
SGPOSGNNEELRYFVRMNRREGIDNNFTGSAVKAASETANMROGQVHAYIRKTG
YDSETEVCNALISMARKGSIISDAEKOPLEVSTKREVSMMAIINYSKRGESLDS
FDOMIHSVRNHNVTIVGLVSAGSHIGLVGDIAGFAFESMSFGLSPEHYVCYDM
LTRAGLSRAKEPTQEMPIKFDALYWRLLSACVYHAKMEIGERFAHHILLEPDSA
TYVLISNLVYASKMDARDLTRQMKKEGVKPEQOSIEYKNSIHSYVGDQNHPLA
DEIHETFDLIRASIEIGYVDCFSLNEIHEODPIIFTHSEKLAISFGILSPAT
VPINVMKMLRVNCNDCHAMIKFVSKVSNREIIVRDAYRPHHEGACSCXDY"
complement(10627. .13689)
/gene="F18A5.40"
/number=1
complement(13690. .13811)
/gene="F18A5.40"
/number=1
complement(13812. .13823)
/gene="F18A5.40"
/number=2
complement(13823. .13823)
/gene="F18A5.50"
complement(join(17299. .17646,18255. .18495,18615. .18818,19108. .19268)
/gene="F18A5.50"
/note="strong similarity to isoflavone reductase-like
protein, lupinus albus, gb:U48590
contains EST gb:T21063, T88563, H36393, AA720215, T76609,
H36674"
/codon_start=1
/product="isoflavone reductase-like protein"
```

```
/protein_id="CAB36830.1"
/db_xref="GI:4455295"
Query Match 28.8%; Score 302.2; DB 8; Length 118718;
Best Local Similarity 97.5%; Pred. No. 3,6e-67;
Matches 307; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 2 CCAAGCATTAATTTTATTTTGTGTCACTAAAGAAAGTCAATGCTCAATTTGAT 61
|||||
DB 96598 CCAAGCATTAATTTTATTTTGTGTCACTAAAGAAAGTCAATGCTCAATTTGAT 96539
62 CGACGACGCTTTTACAGAGAGAGCTTCTGCTATCGACGACATCGAAGCTTCTACAA 121
|||||
DB 96538 CGACGACGCTTTTACAGAGAGAGCTTCTGCTATCGACGACATCGAAGCTTCTACAA 96479
122 TTTCTCCGCTTCTTCTTCTTCTTCTTCTGCTGCTGCTCGACCGTACAACTTCA 181
|||||
DB 96478 TTTCTCCGCTTCTTCTTCTTCTTCTTCTGCTGCTGCTCGACCGTACAACTTCA 96419
182 CGTCCATGGCCAGAGAGAGATCCAAATCCCAATATCCCATATCCGCGCAATTTGCC 241
|||||
DB 96418 CGTCCATGGCCAGAGAGAGATCCAAATCCCAATATCCCATATCCGCGCAATTTGCC 96359
242 TCGTTCACATCACTTCTTCTACATCTATAAGATTTCTCTCCGCTGCGAGCTAG 301
|||||
DB 96358 TCGTTCACATCACTTCTTCTACATCTATAAGATTTCTCTCCGCTGCGAGCTAG 96299
QY 302 GAATTTTCAGCAAT 316
|||||
DB 96298 GAATTTTCAGCAAT 96284
RESULT 5
ATCHRIV37/c
LOCUS ATCHRIV37 199667 bp DNA linear PLN 16-MAR-2000
DEFINITION Arabidopsis thaliana DNA chromosome 4, contig fragment No. 37.
ACCESSION AL161537
VERSION AL161537.2 GI:7268064
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana.
Arabidopsis thaliana.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 21093 to 139810)
Weber,N., Grueninger,D., Schmidheini,T., Mewes,H.W., Lemcke,K. and
Mayer,K.F.X.
Unpublished
2 (bases 1 to 199667)
EU Arabidopsis sequencing, project.
Direct Submission
Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer
Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
lemcke@mips.biochem.mpg.de,mayer@mips.biochem.mpg.de/proj/thai/
Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
E-mail: michael.bevan@bsrc.ac.uk
Information on performance of analysis and a more detailed
annotation of this entry and other sequences of chromosomes 3, 4
and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thai/
this fragment has an overlap with ATCHRIV36 at the 5' end and an
overlap with ATCHRIV38 at the 3' end.
FEATURES
source
location/Qualifiers
1..199667
/organism="Arabidopsis thaliana"
/variety="Columbia"
/db_xref="taxon:3702"
/chromosome="4"
7130..13338
/gene="AT4g13610"
join(7130. .7298,7473. .8347,8449. .8516,9309. .10812,
10929. .11093,11197. .11380,11465. .11576,11636. .11816,
11901. .12098,12231. .12509,12633. .12761,12846. .13061,
```

```

13205..13339)
/gene="At4g13610"
/note="strong similarity to DNA
(cytosine-5-)-methyltransferase, Arabidopsis thaliana ,
PIR2:S59604
Contains C-5 cytosine-specific DNA methylases signatures
AA1077-1089:C-5 cytosine-specific DNA methylases
signatures AA1377-1395
contains EST gb:N38410, A1994691.1, R90103, AA605459"
/codon_start=-1
/product="DNA (cytosine-5-)-methyltransferase-like
protein"
/protein_id="CAB78403.1"
/db_xref="GI:7268065"
/translation="MRTKAGKQKRRSVSDDDVSRERPRKATSGTNFKKSLRSEK
YETVEAKKEQIVDDKEEKEGVFQSFQSVRENTTISQYEDGSPVINTSTVYADCK
PSKKYKKLYDFEKKACAYECVKNLSTNPDTSLKELAAVVRSMNGRKIESSGAVIO
EYVTSQGEFYINOLAGIDERSKHEIKFYNDRLVYSRDSRKTIHKAFSVALRIDES
KVLTSQQLMDGDEDEDKAKLLQEEBHKSMORSRKRSSTTSAPKFTIKINEDI
AHDPADLPSYKNTKDETDLVLFNAGYAVARNLPCTLHNMALVNSDMLISLELP
MKPCADIDVYIQLQIKEMKIDFGEDMIFVLLRTDMAYRLGKPEQYAPWEEILKTV
RIGTSILALIKNETRMAKLSYTVIRKLCGLENDQAYISTFEVRYIVHGOIL
QFLTECPDEYIKRCPVYTGASMODRHHKWTIKKKRKLQKGENLNRKAPKVS
KKMAQATTTLRLNRTMGEPYSYSPDEPLEETGAEPEPEVEDVEEDNEEDTIO
KATVOKADTLKKIRGCKMEIRMBEITIGTCAGEPYLGQALVGRKDVGGAVTL
EVDQGETPLIVETVWFESSDNSKLLHGKLLDRGSTVLTGTANERELFLTNECLV
QLNDIGYVFEIRSRPWHQYKRHEMAADKLDRABEERKADLPLEYCKSLYSPE
KGGFESLPRSDMGLSGFCSSCKIRENEERSKTKLNDSTRELNGIKYSVDGVFO
IPNYLSKDRKRRPVRKYGNVGLRAVVCQIIDYDLKEPKGNTSFYKVRFR
PDDVSAEAYASDIOEYVYSEDYIILPREAIKRCKEVMKKTMDPLCREYVILHVVPC
DRYDSSNGCLKTLPYNNMKESTIKDDTLRKKTETGSAMLIKPDYVKRKLATL
DIRAGCGGLSYGLEKAGVSDTKNAITEEPRAAQFKNHPKTVFVDCNVLLRISML
RLINDRAIMEKCGVDVDCISTTEAAELAKRLDENCKSTLPLPGQVDFISGGPCQOF
SLRNRSDSGMSKNOCOMILAFISFADYFRPKYLENVTFFVFNFGHFIHLTVSL
LENGYOVREGLLEAGAYGISQPRKRAFIAMAAENVEVPEWEPNHWVNGCFKIPLSQ
GLHYAAVOSTKFGAPFRESITVRDAIGDLPIEGESKINKEEMRGSMTVLTDLCKKM
NELNLIRCKIKPTPGADWRDLDEHYVNLNGYVKNRIVPRLKAKDHMKKXKLGTL
DWHGNIPTCTTNTQPMGLVCMCFHPQDRIITISRECARSGFPDSYKFSGNIDKHKQ
VGNVFPPLAFALGRKRLKALHLNLT
7130..7298
/gene="At4g13610"
/number=1
/number=1
7299..7472
/gene="At4g13610"
/number=1
7473..8347
/gene="At4g13610"
/number=2
8348..8448
/gene="At4g13610"
/number=2
8449..8516
/gene="At4g13610"
/number=3
8517..9308
/gene="At4g13610"
/number=3
9309..10812
/gene="At4g13610"
/number=4
10813..10928
/gene="At4g13610"
/number=4
10929..11093
/gene="At4g13610"
/number=5
11094..11196
/gene="At4g13610"
/number=5
11197..11380
/gene="At4g13610"
/number=6
11381..11464

```

```

/gene="At4g13610"
/number=6
11465..11576
/gene="At4g13610"
/number=7
11577..11635
/gene="At4g13610"
/number=7
11636..11816
/gene="At4g13610"
/number=8
11817..11900
/gene="At4g13610"
/number=8
11901..12098
/gene="At4g13610"
/number=9
12099..12230
/gene="At4g13610"
/number=9
12231..12509
/gene="At4g13610"
/number=10
12510..12632
/gene="At4g13610"
/number=10
12633..12761
/gene="At4g13610"
/number=11
12762..12845
/gene="At4g13610"
/number=11
12846..13061
/gene="At4g13610"
/number=12
13062..13204
/gene="At4g13610"
/number=12
13205..13339
/gene="At4g13610"
/number=13
21093..25085
/note="basepairs 1-3993 are not integral part of clone
F1845 but are part ofBAC F723H4"
24246..25412
/gene="At4g13620"
24246..25412
/gene="At4g13620"
/note="similarity to CADMIUM-INDUCED PROTEIN AS30 -
Arabidopsis thaliana, GB:P42736
contains EST gb:AW032262.1, AW035882.1, AW004543.1,
T45770, A1994647.1, T42962, H37693, N65389, A1994192.1,
AA720310, AW029674.1, T76090"
/codon_start=1
/product="putative protein"
/protein_id="CAB78404.1"
/db_xref="GI:7268066"
/translation="MIRPIHTOHLILVINYISPTLSKRLRTGFLIMLTNTQTKNKR
NMEDQPKIETSPMHDKLLSSGIYGFSSSTPOLGVPIFLGMSKSLPASPSTPSY
FVSPHDELTSTHSPVASVPNNLESFSPSOHHPKRPNLTLFLKEPKLIELS
QSESNMSPHYKIIPNSFYQSDONRNEVEINKLTUNTPSGFCNGYSLTKYQPMKSK
TRVVOVTTPTKILRGVRQHRWGKVAEILPRKRTVWLGTETAEQAAMATDTALV
ILRGEFAHLNFPDLKQLKSGSLRCMIASLLEKIOOISSOVSNSSPPPVGTPE
OKNHHMKMESGEDVMMKKQSKREVMGDEVQLSRMPSLMDLIMDALSPHSS"
24246..25412
/gene="At4g13620"
/number=1
26201..28137
/gene="At4g13630"
complement(join(26201..26242,26394..27547,27588..28137))
/gene="At4g13630"
complement(join(26201..26242,26394..27547,27588..28137))
/gene="At4g13630"

```


Db 493 TTATACATAGAGGAACT--TGCAAGCTTCACTAATTCAGTTGTGTTCTTGAG 549
 QY 516 AATTATGATGATTTATGATATTTT--CATTCGTATCCCTCAAGTCTCCACAT 572
 Db 550 AGCAAAATGTTACTGTTCCAGCTTCTTCATGTCAGTTTCCCGAGGATTTAAATG 609
 QY 573 CTTTATGAAGTTTAAACCTTTGAAGTGTGTTGAATTTGTTGACTCTGTGAG 632
 Db 610 TTGCTGAAATAAAGAGATTAAAGAGCTAGAGATTGAAGAGATCACTGAGAA 669
 QY 633 CTTTTCATGACTGTGAGATTAGTATCAAGATGTTGAGATCTTTAGATTAGCAAC 692
 Db 670 CTTTACGTACTGTTGATATCAATTTGAGATTTTGGAGTTGACAGATGTTGCAAT 729
 QY 693 CAAAAAATGTTGAGATTAATAAATGGGCTTCCTCCTACTGAGTGAACCTGTTGC 752
 Db 730 AAAAAGCTGAATATACAGAGACTGAGCCTTAACAGTCTGTTAAACACTCTTAGT 789
 QY 753 AAAGAGCTCTGAAAGCAACAGATCAGCTTGGAACTGGAGTTTATCTCTGTCA 812
 Db 790 AAACAGCTCCTGAAAGCAAGCTATCCGCTGAGCAATTGAGTAATTTCTCTCACT 849
 QY 813 AAGCAGATTTACAAATGAGCAAGATGCTTATGCTTCAATGCAATCTTACAGTT 872
 Db 850 GAGGACCAAGAACTGTATGACGCACTGATGCTTATGCTGTTTATTTATACCAAT 909
 QY 873 CTTAAGGACCTTCTGATGCTGT 895
 Db 910 TTAGAGATTTTGATGATGACTGT 932

RESULT 10

E16596 5189 bp DNA linear PAR 28-JUL-1999
 LOCUS Human WS mRNA, complete cds.
 DEFINITION E16596
 VERSION E16596.1 GI:5711279
 KEYWORDS JP 1998201498-A/1.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.

REFERENCE 1 (bases 1 to 5189)
 AUTHORS Matsunoto, T., Goto, M. and Furuchi, Y.
 TITLE DETECTION OF MUTATION IN PATHOGENIC GENE OF HUMAN WERNER SYNDROME
 JOURNAL Patent: JP 1998201498-A 1 04-AUG-1998;
 EJIJIN KENKYUSHO:KK

COMMENT OS Homo sapiens (human)
 PN JP 1998201498-A/1
 PD 04-AUG-1998
 PF 24-JAN-1997 JP 1997011268
 PI MATSUMOTO TAKEHISA, GOTO MAKOTO, FURUCHI YASUHIRO PC
 C1201/68, C07H21/04, C12N15/09, G01N33/50, G01N33/56; CC
 strandedness: Double;
 CC topology: Linear;
 FH key
 RH Location/Qualifiers

FEATURES
 source 1. .5189
 Location/Qualifiers
 1. .5189
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 CDS 232..4530
 /product="WS"
 /note="Werner's syndrome".
 FT
 FT

BASE COUNT 1663 a 889 c 1129 g 1508 t
 ORIGIN

Query Match 10.6%; Score 111; DB 6; Length 5189;
 Best Local Similarity 54.1%; Pred. No. 8e-18;
 Matches 272; Conservative 0; Mismatches 225; Indels 6; Gaps 2;

QY 396 GATACCAAGAGATGAATCTGGAATGCTTTGCTTGGCTTGATATGACTGAGACCA 455
 Db 433 GATATTTAGCAGACTATCAGATGAGGATTTGGATTTGACATGAGATGGCCACA 492
 QY 456 AGTTTTGAAGAGTGTTCGCCGGGAGGTTCGACCTGCGCATATGATGATGATAGT 515
 Db 493 TTATACATAGAGGAACT--TGCAAGCTTCACTAATTCAGTTGTGTTCTGAG 549
 QY 516 AATTATGATGATTTATGATATTTT--CATTCGTATCCCTCAAGTCTCCACAT 572
 Db 550 AGCAAAATGTTACTGTTCCAGCTTCTTCATGTCAGTTTCCCGAGGATTTAAATG 609
 QY 573 CTTTATGAAGTTTAAACCTTTGAAGTGTGTTGAATTTGTTGACTCTGTGAG 632
 Db 610 TTGCTGAAATAAAGAGATTAAAGAGCAGTGTAGAGATTGAAGAGATCACTGAAA 669
 QY 633 CTTTTCATGACTGTGAGATTAGTATCAAGATTTGAGATCTTTAGATTAGCAAC 692
 Db 670 CTTTACGTGACTTGTGATATCAAAATTTGAAGATTTTGGAGTTGACAGATGTTGCCAAT 729
 QY 693 CAAAAAATGTTGAGATTAATAAATGGGCTTCCTCCTACTGAGTGAACCTGTTGC 752
 Db 730 AAAAAGCTGAATATACAGAGACTGAGCCTTAAACAGTCTGTTAAACACTCTTAGT 789
 QY 753 AAAGAGCTCTGAAAGCAACAGATCAGCTTGGAACTGGAGTTTATCTCTGTCA 812
 Db 790 AAACAGCTCCTGAAAGCAAGCTATCCGCTGAGCAATTGAGTAATTTCTCTCACT 849
 QY 813 AAGCAGATTTACAAATGAGCAAGATGCTTATGCTTCAATGCAATCTTACAGTT 872
 Db 850 GAGGACCAAGAACTGTATGACGCACTGATGCTTATGCTGTTTATTTATACCAAT 909
 QY 873 CTTAAGGACCTTCTGATGCTGT 895
 Db 910 TTAGAGATTTTGATGATGACTGT 932

RESULT 11

HUMDR 5189 bp ss-DNA linear PRI 25-APR-1996
 LOCUS Homo sapiens Werner syndrome gene, complete cds.
 DEFINITION L76937.1 GI:1280207
 ACCESSION L76937.1
 VERSION Werner's syndrome; progeria.
 KEYWORDS Homo sapiens DNA.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 5189)
 AUTHORS Yu, C.-E., Oshima, J., Fu, Y.-H., Wjisman, E.M., Hisama, F., Allisch, R.,
 Matthews, S., Nakura, J., Miki, T., Ouals, S., Martin, G.M., Mulligan, J.,
 and Schellenberg, G.D.
 TITLE Positional cloning of the Werner's syndrome gene
 JOURNAL Science 272 (5259), 258-262 (1996)
 MEDLINE 96181115
 PUBMED 8602509

FEATURES
 source 1. .5189
 Location/Qualifiers
 1. .5189
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 mRNA 1. .5189
 /partial
 CDS 232..4530
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="AAC11981.1"
 /db_xref="GI:1280208"
 /translation="MSEKLTETTAQOKCEPMANVONKCAVEERRACVRSVFEDDL
 PLEFTGSIYSDASDCSFLSEBIDISMSLSDGVGVGDMPPLYNKGKLGKVALIOL
 CVSESKCYLFHVSMSVFPQGLMKLEENKRAVKAAGVIEGDOKLILFEDIKLNFYE
 LTVANKKCTETWMSLVKHLGLKOLKDSIRCSNNSKFPVFDOKLVATDAY
 AGPIYANKEILDDYVORPAINKKEEFLISDMKOLTSISEYMDLAKHLPAPSKIE
 NPNRAVSLILKDISBNLISLRMIIGSTINITEELRPSNNNLISFEDSTGGVQOKQIR

HEVLIIHVEDETWPTLHLAKHDEVLGNKVERKEDGFEDVEDNKLKENMERACI
MSLDITHELIQLEOQSEYLSIDIAKSTELSPNDENDSIYSESEDJEMELK
HSPNDENDSIYSESEDJEMELKSTENLNSGVEPDKKMEINDELPRKE
BEDDENANEGEDDDKDFLMPNBEQVTCMKFGHSEFPYOMKVYHSLBERD
NVAVMATGKSLCFQYPPVYKIGLVISPLISMEDOVLDAKSNIPACLGSAQS
EVALDIDKIGKRIYVTPPEYSGNGLQOEADIGITLIVADAHCSIEGHDFRD
SFRKLSLKTALPMPVIALPATPASSIREDIVRCINLNPOTITCGEPRLMIEVR
RKTGNILQLOPFLVTSNMEFEGPTIICPSRKMTQOVTGELKRLNISCCTYHAE
SESTRKDIHREVRDEIOCVIATIFAGMINKADROYHYCAPDMESYOEIRAG
RDGLOSCCHVLPADINLRHLTEIRNEKRLIKLMAKMEKYLHSSRRROIIL
SHFEDKOVOKASLGIMTEKCCNCRSLDHCTSMDSSEDITSMDGPQAFKLLSADI
LGEKFGIGLPIFLRGSNSQRLADQYRRHSLEGTGKDOYESWMAKFSROLITGEGLVE
VSRYNKFMKICALTRKGRNMLKRNKATESQSILQANEEICPKFLPSKTSYSGTKE
HCYNOVPELSTEKSNLEKTSYKPCDISSGNSIKSINVOSEPEKAYSSOPVIS
AOEOTOIYLKGLVEAROKHANKMDVPAIATKILYDMAKMRPTIYENKRIDYV
SEGRKAMLPALLEVIRKHCOTNSVOTDLSSTKPOEOTSTIVANKITCTLSQNAIT
YSLPOEKMRPLKSLIESRLPLMTIGMHSQVAKGCPDLERAGJLEPVOKITADVI
RNPVNSDKSISLRMLVPENIDTYLIHMALEILKHGDSGLPSDCVKNKRCPEGS
EICSSSKRSKEEVIINTETSAERKRRLPVPFAKGSPTSKLMDKTRKGLFS"

polyA_signal
polyA_site
BASE COUNT 1663 a 889 c 1129 g 1508 t
ORIGIN

Query Match 10.6%; Score 111; DB 9; Length 5189;
Best Local Similarity 54.1%; Pred. No. 8e-18;
Matches 272; Conservative 0; Mismatches 225; Indels 6; Gaps 2;

QY 396 GATACCAAGAGAGATGATCTGGAATAGCTTTGTTGGCTTGATATTAGTGAGACCA 455
||||| - - - - - ||||| - - - - - ||||| - - - - - ||||| - - - - - |||||
DB 433 GATATTAGCATGAGCTATACAGATGGGAGTGTGGTGAGTTGACATGAGGCGACCA 492
456 AGTTTAAAGAAAGTGTCTCCCGGGGAAAGTTGGCAGCTGCAGATATGTAGATAGT 515
- - - - - ||||| - - - - - ||||| - - - - - ||||| - - - - - |||||
DB 493 TTATACAAATAGAGGAAACT---TGCCAAAGTTGCATATTACGTTGTGTCTTGAG 549
- - - - - ||||| - - - - - ||||| - - - - - ||||| - - - - - |||||
QY 516 AATTATTGTGATGTATGCAATATTTT---CATTCGTGATCCCTCAAGTCTCCAAACAT 572
- - - - - ||||| - - - - - ||||| - - - - - ||||| - - - - - |||||
DB 550 AGCAAAATGTTACTTGTCCACGTTCTTCATGATGATGTTTCCAGAGGATTAAGG 609
- - - - - ||||| - - - - - ||||| - - - - - ||||| - - - - - |||||
QY 573 CTATTGAAGATTCAACACTTTAAAGTAGTATTGCAATTGATGCTGCTGTGAG 632
- - - - - ||||| - - - - - ||||| - - - - - ||||| - - - - - |||||
DB 610 TTGCTGAAATAAAGCACTTAAAGGAGGTGAGAAATGAAGAGATCAGTGAAA 669
- - - - - ||||| - - - - - ||||| - - - - - ||||| - - - - - |||||
QY 633 CTTTCCATGACTAGTATGATCAAGAGATGTTGAGATCTTGACATTAGGCAAC 692
||||| - - - - - ||||| - - - - - ||||| - - - - - ||||| - - - - - |||||
DB 670 CTTCACGTGACTTTGATATCAAAATGAAGAAATTTGTGAGTTGACAGATGTGCCAAT 729
693 CAAAAAATTTGTGAGATAAAAAATGGGCTTGGCTCCTACATACTGACACTGTTTGC 752
||||| - - - - - ||||| - - - - - ||||| - - - - - ||||| - - - - - |||||
DB 730 AAAAAGCTGAATATGACAGAGACCTTGACAGCTTGATTAACACCTCTTAGT 789
733 AAAGAGCTCTGACCAACAAAGATCAGGCTTGGGAACTGGAGTTTATCTCTGTGCA 812
||||| - - - - - ||||| - - - - - ||||| - - - - - ||||| - - - - - |||||
DB 790 AAACAGCTCTGAAAAGCAAGTCTATCCGCTAGCAATGGAGTAAATTTCTCTCACT 849
813 AAGCAGAGCTTACGACAGCAAGGATGCTAATGCTTANTGCGATCTTACAAAGT 872
850 GAGGACCAAGAACTGATGACAGCACTGATGCTTATGCTGTTTATTTATTTACGAAAT 909
QY 873 CTTAAGAGACTTCTGATGCTGT 895
- - - - - ||||| - - - - - ||||| - - - - - ||||| - - - - - |||||
DB 910 TTAGAGATTTTGATGATGATCTGT 932

RESULT 12
AF091214 5208 bp mRNA linear pri 08-OCT-1998
LOCUS AF091214
DEFINITION Homo sapiens WRN (WRN) mRNA, complete cds.
ACCESSION AF091214
VERSION AF091214.1 GI:3719420
KEYWORDS
SOURCE Homo sapiens.

ORGANISM

Homo sapiens

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS

1 (bases 1 to 5208)
Paepker,B.W., Gayle,M., Brady,W., Swartz,A., Gillett,L.A.,

TITLE

Alisch,R.S., Mulligan,J., Galas,D. and Fu,Y.-H.
Genomic structure of the human Werner's gene and cloning of the

JOURNAL

mouse homology
Unpublished

AUTHORS

2 (bases 1 to 5208)
Paepker,B.W., Gayle,M., Brady,W., Swartz,A., Gillett,L.A.,

TITLE

Alisch,R.S., Mulligan,J., Galas,D. and Fu,Y.-H.
Direct Submission

JOURNAL

Submitted (12-SEP-1998) Genomics, Chiroscience R&D Inc., 1631 220th
Street SE, Bothell, WA 98021, USA

FEATURES

Location/Qualifiers
1..5208
/organism="Homo sapiens"

gene

1..5208
/db_xref="taxon:9606"

5'UTR

1..231
/gene="WRN"

CDS

232..4530
/gene="WRN"

ORIGIN

/note="Werner syndrome"

ORIGIN

/product="WRN"

ORIGIN

/protein_id="AAC63361.1"

ORIGIN

/db_xref="GI:3719421"

ORIGIN

/translation="MSEKLETTAAQRCPEMMNYYONKCAVEERACVRSVFEDDL
PLEFTGSIYVSADSDCSFLESDISMSDDDVGVGDEMPLVNRGLGVALIOL
CYSESKCYLFHVSMSVFPQIGKMLLENKAYKAGVGTGMDKMLTRDITLKNVE
LDVANKRKICTEYMSLNSLVHILGKOLSKRSICRSMSRPLTEDOKIATADY
AGTITRNLIELDDVYORFANKNEEILISDMKOLTSISEVMIAKRLPAFSKLE
NPRVSLIKLDISENL.SLRMIIGSTNLETLPBNNILNLFEDSTGVOQOIR
EHEVLIIHVEDETWPTLHLAKHDEVLGNKVERKEDGFEDVEDNKLKENMERACI
MSLDITHELIQLEOQSEYLSIDIAKSTELSPNDENDSIYSESEDJEMELK
HSPNDENDSIYSESEDJEMELKSTENLNSGVEPDKKMEINDELPRKE
BEDDENANEGEDDDKDFLMPNBEQVTCMKFGHSEFPYOMKVYHSLBERD
NVAVMATGKSLCFQYPPVYKIGLVISPLISMEDOVLDAKSNIPACLGSAQS
EVALDIDKIGKRIYVTPPEYSGNGLQOEADIGITLIVADAHCSIEGHDFRD
SFRKLSLKTALPMPVIALPATPASSIREDIVRCINLNPOTITCGEPRLMIEVR
RKTGNILQLOPFLVTSNMEFEGPTIICPSRKMTQOVTGELKRLNISCCTYHAE
SESTRKDIHREVRDEIOCVIATIFAGMINKADROYHYCAPDMESYOEIRAG
RDGLOSCCHVLPADINLRHLTEIRNEKRLIKLMAKMEKYLHSSRRROIIL
SHFEDKOVOKASLGIMTEKCCNCRSLDHCTSMDSSEDITSMDGPQAFKLLSADI
LGEKFGIGLPIFLRGSNSQRLADQYRRHSLEGTGKDOYESWMAKFSROLITGEGLVE
VSRYNKFMKICALTRKGRNMLKRNKATESQSILQANEEICPKFLPSKTSYSGTKE
HCYNOVPELSTEKSNLEKTSYKPCDISSGNSIKSINVOSEPEKAYSSOPVIS
AOEOTOIYLKGLVEAROKHANKMDVPAIATKILYDMAKMRPTIYENKRIDYV
SEGRKAMLPALLEVIRKHCOTNSVOTDLSSTKPOEOTSTIVANKITCTLSQNAIT
YSLPOEKMRPLKSLIESRLPLMTIGMHSQVAKGCPDLERAGJLEPVOKITADVI
RNPVNSDKSISLRMLVPENIDTYLIHMALEILKHGDSGLPSDCVKNKRCPEGS
EICSSSKRSKEEVIINTETSAERKRRLPVPFAKGSPTSKLMDKTRKGLFS"

ORIGIN

4531..5208
/gene="WRN"

ORIGIN

3'UTR

ORIGIN

BASE COUNT 1681 a 887 c 1131 g 1509 t

ORIGIN

Query Match 10.6%; Score 111; DB 9; Length 5208;
Best Local Similarity 54.1%; Pred. No. 8e-18;
Matches 272; Conservative 0; Mismatches 225; Indels 6; Gaps 2;

QY 396 GATACCAAGAGAGATGATCTGGAATAGCTTTGTTGGCTTGATATTAGTGAGACCA 455
||||| - - - - - ||||| - - - - - ||||| - - - - - ||||| - - - - - |||||
DB 433 GATATTAGCATGAGCTATACAGATGGGAGTGTGGTGAGTTGACATGAGGCGACCA 492
456 AGTTTAAAGAAAGTGTCTCCCGGGGAAAGTTGGCAGCTGCAGATATGTAGATAGT 515
- - - - - ||||| - - - - - ||||| - - - - - ||||| - - - - - |||||
DB 493 TTATACAAATAGAGGAAACT---TGCCAAAGTTGCATATTACGTTGTGTCTTGAG 549
- - - - - ||||| - - - - - ||||| - - - - - ||||| - - - - - |||||
QY 516 AATTATTGTGATGTATGCAATATTTT---CATTCGTGATCCCTCAAGTCTCCAAACAT 572

Db 550 AGCAATGTTACTGTTCCACAGTCTTCCATGTCACATTTTCCCGAGGATTAATAATG 609
QY 573 CTTATTAAGATTCACACACTTGTAAAGTAGTATGTAATGATGCTGCTGTAAG 632
Db 610 TTGCTTAAATTAAGACAGTTAAAGAGGAGTGTAGAAATGAGAGTACAGTGA 669
QY 633 CTTTTCATGACTATGAGTGTATGTAAGATGTTAGAGATCTTCAATTAAGCCAC 692
Db 670 CTTTACGTCATGTTGATATCAATTAAGAAATTTTGTGAGTGTAGACAGATGTTCCAA 729
QY 693 CAAAATTTGCTGAGATTAATAATGAGGCTTCCCTCACTACTGACACACTGTTTC 752
Db 730 AAAAGGCTGAATGTACAGAGACCTGAGCTTAAACAGTCTGTTAAACACTTGTAGT 789
QY 753 AAAGAGCTCTGAAGCCAAACAGATCAGCTTGGAGATGGAGATTTTATCCTCTGCA 812
Db 790 AAACAGCTCTGAAGACAAAGTCTATCCGCTGTAGCAATGAGATTAATTTCTCTACT 849
QY 813 AACGACGATTTACATACGACGACAAAGATGCTTATGCTTATGATGATCTTTACAGGTT 872
Db 850 GAGGACCGAAGAACTGTATGACGACGATGATGCTTATGCTTATTTATTTACCAAT 909
QY 873 CTTAAGACCTTCTGATGCTGT 895
Db 910 TTAGAGATTTGATGATGATGCT 932

RESULT 13
AF067418 5143 bp mRNA linear VRT 09-OCT-1998

LOCUS Xenopus laevis focus forming activity 1 (FFA-1) mRNA, complete cds.
DEFINITION AF067418
ACCESSION AF067418.1 GI:3420290
VERSION
KEYWORDS
SOURCE
ORGANISM

Xenopus laevis.
Xenopus laevis.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;

Xenopus.
Xenopus laevis.
1 (bases 1 to 5143)
Yan, H., Chen, C. Y., Kobayashi, R. and Newport, J.
Replication focus-forming activity 1 and the Werner syndrome gene

product
Nat. Genet. 19 (4), 375-378 (1998)
JOURNAL
MEDLINE
PUBMED
2 (bases 1 to 5143)
9697700
REFERENCE
AUTHORS
TITLE
JOURNAL

Submitted (20-May-1998) Fox Chase Cancer Center, 7701 Burholme
Avenue, Philadelphia, PA 19111, USA
Location/Qualifiers
1. 5143
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/cell_type="oocyte"
1. 5143
/gene="FFA-1"
160. 4470
/gene="FFA-1"
/function="DNA helicase"
/codon_start=1
/product="focus forming activity 1"
/protein_id="AAC63512.1"
/translation="MTSLQKLPENMSVYQOEDRIDAKSPCKKNILENLPKRN
GSIIVSYESWDCSLSDIRSSLEEDVLEFDLEPVTYKGTGKVALIQVCSEK
CYLPHISPMAGPFGKRLLEDSVRKVGIGIDDKMLSDYELKIGFIELSEMAN
OKLRKEKTPFNGLIKHLFKELQYKRSYSGSMWDFLLEDDKLYATATYAGLLY
KLEGMADHESDSFRVREGVADCKGVKRLDTLSGLMDLVNOVPSFCYTEAHVA
VDLEDESEKLEELRNIMKASKAGNGLHFGNSDPSCKDKSLIHVACKESLAEHNM
DCKNADSONKNDIDSCONENRDEDFMTLGISEBELYMEREDDKOTNPDIYLNKDS

FEATURES
source
gene
CDS

CDTNEKMSVIESDEDPSEIILKSLIEDLNSTEEALGTGPQAGLIPAKSVDTYAD
LVEEAGIEEDDDDDMDPMSPEPAOHISOLKTYFGSSSRKPOMKVHSLERBN
LYEVAATGKSLCYOFAPVYSGIYICPLISMEQVOLFESNTSCFLSAGSK
NLOADVADGKMRVYVMPEFCRISILLOLDNRYGTLIAIDAHISMGHRSRA
YRISGSLKRLHPNVPYALTRATSPSIREDTISLHNPQVCTSDRNILYDAR
KTNISTDLQFLTKOQSGMEFEGATYYCPKRTSEVTAELILUGIACGYHAG
MGIKQREHHRHREMRDEHCVAVAVAGMGINRDIKRVHIGAPKMESEYQOEGA
GRDLPSCCHALMAQADNMNRHMLGETIPKGFREYILKMLTKMEKYLNSGCRKI
LSHEDROLKRASSGIMGTCEKDCNCRILCINSINDIEDNLODFGPQAKFTSAD
VLQKFGTGVPLVLETRGTSORVDRPDRNHSLSGSDQTEAFMKVLABRLITEGYO
ESSQOTFPTICLISKSNMLIKANNECCPSILIPSNELCLOFRTVSPSSQAS
SNVPHASNTRSSMPKAGPERKMDKDFSTQEBERLSKAGVSKSRKLPCKLSNP
PEPEVSPRELEDTTLVGRVVAROKIASERDILPAVALANKVAVDAKLPTESEM
KILQVSEAKSAMLAPLEEVKERCINLSNDVIFSQVQSESTFPTPREOERISLP
ESORMSVLFOEONLSKTIADVCLNAVYGMOLAKAGSFDVQAGLTPENMK
LITYAIKPPINSDFSKAIREVYPAINDIGYPRVATYILEKTKSSGALQPIPMAS
KTLIOTEPENKNSVONTKRVYMGKSMLEKPPKPTQATYELTEVGGKATPIMLAS
NMSLDQTEBELSESOSTTRPRRRLPENFGSTKGNATRCIOESKNLEBEGSPFD

BASE COUNT 1689 a 951 c 1104 g 1399 t
ORIGIN
Query Match 10.5% Score 110.6; DB 5; Length 5143;
Best Local Similarity 54.1% Pred. No. 1e-17;
Matches 271; Conservative 0; Mismatches 224; Indels 6; Gaps 2;

QY 423 GCTTTGTGGCTGATATTTAGTGAAGACCAAGTTTGAAGAGGTTCTCCGGGG 482
Db 370 GATGTACTTGTTTGTATATAGATGCGCTTCTTATACAAAGG---AAAACGGA 426
QY 483 AAGGTGCACTGTCAGATATGTGTAGATGATATTTATTTGATATGATATTTT 542
Db 427 AAAGTCTCTTATACAGTGTGTGTCTGAGAAAAAGCTATTTGTTCAATATTTCA 486
QY 543 CAT---TCTGATACCTCAAGCTCCCAACATCTTATTAAGATTAACACTGTAAAG 599
Db 487 CTTATGCTGCTTCTCTTAAGGTTTGAAGAGGCTGTAGAGATGAGTCTGTAGAAA 546
QY 600 GTAGATATGGAATTTGATGTGACTCTGTGAAGCTTTTCAATGATGAGTATGATC 659
Db 547 GTTGCTGTGGGAATTTGAAGGACCACTGGAAGTATTTGATGAGTCAAGCTTAACTT 606
QY 660 AAATATTTGAGATCTTTCAAGATTTAGCCAAACCAAAATTTGATGAGTAAATG 719
Db 607 AAAGCTTTATTTGACCTTGGGAATGCTAACCAAAACATACGCTCAAGGAAATGG 666
QY 720 GACCTTGCCTACATGAGACACTGTTTGAAGAGCTCTGAGACCAACAGATC 779
Db 667 ACCTTTAATGTTGATCAAAACATCTTTCAAGAGCAACTTTATTAAGAAAGCTTAC 726
QY 780 AGGCTTGGAACTGGAGTTTATCTCTGTCAAAAGCAGCATTAATACGACAGC 839
Db 727 CGTTGACAGCAACTGGATATTTTCTCTCAGTCAAGAGAAATTTGTCGCCACT 786
QY 840 GATGCTTATGCTTATGATGATCTTTTACAGAGTTTAAAGACCTCTGATGCTGAGT 899
Db 787 GATGCTTATGCTGAGCTTATTTTACAGAACTAAGAGGAGCATGATGCTCAGAGAGT 846
QY 900 GGCTCATACGATGAGAGGA 920
Db 847 GATTCCTTCCTGCTGTTGCACA 867

RESULT 14
ATPFA0/c 200576 bp DNA linear PLN 28-JUN-1999
DEFINITION Arabidopsis thaliana DNA chromosome 4, ESSA I FCA contig fragment
ACCESSION No. 0.
VERSION 297335
KEYWORDS 297335.2 GI:5280985
SOURCE Arabidopsis thaliana.
ORGANISM Arabidopsis thaliana

REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
AUTHORS	1 (bases 1 to 200576) Bevan,M., Stiekema,W., Murphy,G., Wandut,R., Pohl,T., Terryn,N., Kreis,M., Kavanagh,T., Entlan,K.D., Rieger,M., James,R., Puidomenech,P., Hatzopoulos,P., Obermayer,B., Duesterhoft,A., Jones,J., Palme,K., Ansoorge,W., Dalseny,M., Bancroft,I., Mewes,H.W., Schueller,C. and Chalmatzis,N.		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 200576) EU Arabidopsis sequencing project.		
AUTHORS	Direct Submission		
JOURNAL	Submitted (25-JUN-1999) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: schnell@mpi-biochem.mpg.de, mayer@mpi-biochem.mpg.de Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bsrc.ac.uk		
COMMENT	On Jun 30, 1999 this sequence version replaced gi:2244747. Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/ this fragment has an overlap with AFCA1 at the 3' end.		
FEATURES	Location/Qualifiers		
source	1..200576 /organism="Arabidopsis thaliana" /variety="Columbia" /db_xref="taxon:3702" /chromosome="4" 1..63289 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM020K8" 62805..140141 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM06G17" 134184..200576 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM08H22" 1753..4428 /gene="d13000w" join(1753..3194,3261..4428) /gene="d13000w" /note="similarity to Cf-2.2, Solanum pimpinellifolium" /codon_start=1 /product="disease resistance Cf-2 like protein" /protein_id="CAB10171.1" /db_xref="GI:2244748" /db_xref="SPTREMBL:O23253" /translation="MTITWISLCLIFCNSNLIYIAKDLCPDORDALFEKNESIP SPDSDLMLIQTAKARNNTDCCSWGIGSDPRTGVVVELDGNLGRNSSLF RLQHLQSIDLSTNDLSTLPDSSGNFKYRLVNLGLNFGELPTSLNLSLYLDDL SYNDLITGEIYDLSNGNLKRLRYLSLFCSEKPKIPSSICNLITLIDDLISMYTFEEL PDSKGNLSIRLVNLRHCRNFEGKIPSTGSLNLDLISKNEFTSEGDSSSLRL TDFQMLNLNLVNDLSSNQFKALPSPNMSLSLLEAFDISGNSFCSTIPSSLEML PSLIRKDLGTNDPSPGLKIGNISPSNLOELYIGENNINGPIPRSTLKLVGSAISL FMDTGIVDFISFQILKSLRSLDLSGINLISSSHLPSMMHLILSCNISOFPKEL ENQTSILYHLDISANOIEGOVPEMLRPLTSLFASDNKPSGEIPRAVCEIGTLVLSNN NFGSIPPCFEISNKTSLTSLIRNNSLSGVIPESLHGLIRSLIDVGSNLSQFPKSL INCSTIQLPLANTENKRIINDTFPSMLKSLPMLQVLVLRKSNFHGPITPSGSLSPKRL FDISNRSGLVPSDYFVGWVSWSLVDIIDMTPGFTVVGDOESFKSVLTIKGLN MELVSGEPIYKTIIVSGNRLBGDIPESIGIKELIVLMSNNAFGH1PSPISNLSN IQSLDLSQNRSLSGIPGELGELFTLARMFYSNMEGPIPOGTIOIOSNSSFAENPG LCGAPLKKKCGEEEDKEKKDKGLSWVAAGVYVPELFGALIGHILTSYKRWME WRFSCFSPLKK" /number=1 /gene="d13000w" 1753..3194 /number=1 3195..3260 /gene="d13000w"		
source	1..63289 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM020K8" 62805..140141 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM06G17" 134184..200576 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM08H22" 1753..4428 /gene="d13000w" join(1753..3194,3261..4428) /gene="d13000w" /note="similarity to Cf-2.2, Solanum pimpinellifolium" /codon_start=1 /product="disease resistance Cf-2 like protein" /protein_id="CAB10171.1" /db_xref="GI:2244748" /db_xref="SPTREMBL:O23253" /translation="MTITWISLCLIFCNSNLIYIAKDLCPDORDALFEKNESIP SPDSDLMLIQTAKARNNTDCCSWGIGSDPRTGVVVELDGNLGRNSSLF RLQHLQSIDLSTNDLSTLPDSSGNFKYRLVNLGLNFGELPTSLNLSLYLDDL SYNDLITGEIYDLSNGNLKRLRYLSLFCSEKPKIPSSICNLITLIDDLISMYTFEEL PDSKGNLSIRLVNLRHCRNFEGKIPSTGSLNLDLISKNEFTSEGDSSSLRL TDFQMLNLNLVNDLSSNQFKALPSPNMSLSLLEAFDISGNSFCSTIPSSLEML PSLIRKDLGTNDPSPGLKIGNISPSNLOELYIGENNINGPIPRSTLKLVGSAISL FMDTGIVDFISFQILKSLRSLDLSGINLISSSHLPSMMHLILSCNISOFPKEL ENQTSILYHLDISANOIEGOVPEMLRPLTSLFASDNKPSGEIPRAVCEIGTLVLSNN NFGSIPPCFEISNKTSLTSLIRNNSLSGVIPESLHGLIRSLIDVGSNLSQFPKSL INCSTIQLPLANTENKRIINDTFPSMLKSLPMLQVLVLRKSNFHGPITPSGSLSPKRL FDISNRSGLVPSDYFVGWVSWSLVDIIDMTPGFTVVGDOESFKSVLTIKGLN MELVSGEPIYKTIIVSGNRLBGDIPESIGIKELIVLMSNNAFGH1PSPISNLSN IQSLDLSQNRSLSGIPGELGELFTLARMFYSNMEGPIPOGTIOIOSNSSFAENPG LCGAPLKKKCGEEEDKEKKDKGLSWVAAGVYVPELFGALIGHILTSYKRWME WRFSCFSPLKK" /number=1 /gene="d13000w" 1753..3194 /number=1 3195..3260 /gene="d13000w"		
source	1..63289 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM020K8" 62805..140141 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM06G17" 134184..200576 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM08H22" 1753..4428 /gene="d13000w" join(1753..3194,3261..4428) /gene="d13000w" /note="similarity to Cf-2.2, Solanum pimpinellifolium" /codon_start=1 /product="disease resistance Cf-2 like protein" /protein_id="CAB10171.1" /db_xref="GI:2244748" /db_xref="SPTREMBL:O23253" /translation="MTITWISLCLIFCNSNLIYIAKDLCPDORDALFEKNESIP SPDSDLMLIQTAKARNNTDCCSWGIGSDPRTGVVVELDGNLGRNSSLF RLQHLQSIDLSTNDLSTLPDSSGNFKYRLVNLGLNFGELPTSLNLSLYLDDL SYNDLITGEIYDLSNGNLKRLRYLSLFCSEKPKIPSSICNLITLIDDLISMYTFEEL PDSKGNLSIRLVNLRHCRNFEGKIPSTGSLNLDLISKNEFTSEGDSSSLRL TDFQMLNLNLVNDLSSNQFKALPSPNMSLSLLEAFDISGNSFCSTIPSSLEML PSLIRKDLGTNDPSPGLKIGNISPSNLOELYIGENNINGPIPRSTLKLVGSAISL FMDTGIVDFISFQILKSLRSLDLSGINLISSSHLPSMMHLILSCNISOFPKEL ENQTSILYHLDISANOIEGOVPEMLRPLTSLFASDNKPSGEIPRAVCEIGTLVLSNN NFGSIPPCFEISNKTSLTSLIRNNSLSGVIPESLHGLIRSLIDVGSNLSQFPKSL INCSTIQLPLANTENKRIINDTFPSMLKSLPMLQVLVLRKSNFHGPITPSGSLSPKRL FDISNRSGLVPSDYFVGWVSWSLVDIIDMTPGFTVVGDOESFKSVLTIKGLN MELVSGEPIYKTIIVSGNRLBGDIPESIGIKELIVLMSNNAFGH1PSPISNLSN IQSLDLSQNRSLSGIPGELGELFTLARMFYSNMEGPIPOGTIOIOSNSSFAENPG LCGAPLKKKCGEEEDKEKKDKGLSWVAAGVYVPELFGALIGHILTSYKRWME WRFSCFSPLKK" /number=1 /gene="d13000w" 1753..3194 /number=1 3195..3260 /gene="d13000w"		
source	1..63289 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM020K8" 62805..140141 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM06G17" 134184..200576 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM08H22" 1753..4428 /gene="d13000w" join(1753..3194,3261..4428) /gene="d13000w" /note="similarity to Cf-2.2, Solanum pimpinellifolium" /codon_start=1 /product="disease resistance Cf-2 like protein" /protein_id="CAB10171.1" /db_xref="GI:2244748" /db_xref="SPTREMBL:O23253" /translation="MTITWISLCLIFCNSNLIYIAKDLCPDORDALFEKNESIP SPDSDLMLIQTAKARNNTDCCSWGIGSDPRTGVVVELDGNLGRNSSLF RLQHLQSIDLSTNDLSTLPDSSGNFKYRLVNLGLNFGELPTSLNLSLYLDDL SYNDLITGEIYDLSNGNLKRLRYLSLFCSEKPKIPSSICNLITLIDDLISMYTFEEL PDSKGNLSIRLVNLRHCRNFEGKIPSTGSLNLDLISKNEFTSEGDSSSLRL TDFQMLNLNLVNDLSSNQFKALPSPNMSLSLLEAFDISGNSFCSTIPSSLEML PSLIRKDLGTNDPSPGLKIGNISPSNLOELYIGENNINGPIPRSTLKLVGSAISL FMDTGIVDFISFQILKSLRSLDLSGINLISSSHLPSMMHLILSCNISOFPKEL ENQTSILYHLDISANOIEGOVPEMLRPLTSLFASDNKPSGEIPRAVCEIGTLVLSNN NFGSIPPCFEISNKTSLTSLIRNNSLSGVIPESLHGLIRSLIDVGSNLSQFPKSL INCSTIQLPLANTENKRIINDTFPSMLKSLPMLQVLVLRKSNFHGPITPSGSLSPKRL FDISNRSGLVPSDYFVGWVSWSLVDIIDMTPGFTVVGDOESFKSVLTIKGLN MELVSGEPIYKTIIVSGNRLBGDIPESIGIKELIVLMSNNAFGH1PSPISNLSN IQSLDLSQNRSLSGIPGELGELFTLARMFYSNMEGPIPOGTIOIOSNSSFAENPG LCGAPLKKKCGEEEDKEKKDKGLSWVAAGVYVPELFGALIGHILTSYKRWME WRFSCFSPLKK" /number=1 /gene="d13000w" 1753..3194 /number=1 3195..3260 /gene="d13000w"		
source	1..63289 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM020K8" 62805..140141 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM06G17" 134184..200576 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM08H22" 1753..4428 /gene="d13000w" join(1753..3194,3261..4428) /gene="d13000w" /note="similarity to Cf-2.2, Solanum pimpinellifolium" /codon_start=1 /product="disease resistance Cf-2 like protein" /protein_id="CAB10171.1" /db_xref="GI:2244748" /db_xref="SPTREMBL:O23253" /translation="MTITWISLCLIFCNSNLIYIAKDLCPDORDALFEKNESIP SPDSDLMLIQTAKARNNTDCCSWGIGSDPRTGVVVELDGNLGRNSSLF RLQHLQSIDLSTNDLSTLPDSSGNFKYRLVNLGLNFGELPTSLNLSLYLDDL SYNDLITGEIYDLSNGNLKRLRYLSLFCSEKPKIPSSICNLITLIDDLISMYTFEEL PDSKGNLSIRLVNLRHCRNFEGKIPSTGSLNLDLISKNEFTSEGDSSSLRL TDFQMLNLNLVNDLSSNQFKALPSPNMSLSLLEAFDISGNSFCSTIPSSLEML PSLIRKDLGTNDPSPGLKIGNISPSNLOELYIGENNINGPIPRSTLKLVGSAISL FMDTGIVDFISFQILKSLRSLDLSGINLISSSHLPSMMHLILSCNISOFPKEL ENQTSILYHLDISANOIEGOVPEMLRPLTSLFASDNKPSGEIPRAVCEIGTLVLSNN NFGSIPPCFEISNKTSLTSLIRNNSLSGVIPESLHGLIRSLIDVGSNLSQFPKSL INCSTIQLPLANTENKRIINDTFPSMLKSLPMLQVLVLRKSNFHGPITPSGSLSPKRL FDISNRSGLVPSDYFVGWVSWSLVDIIDMTPGFTVVGDOESFKSVLTIKGLN MELVSGEPIYKTIIVSGNRLBGDIPESIGIKELIVLMSNNAFGH1PSPISNLSN IQSLDLSQNRSLSGIPGELGELFTLARMFYSNMEGPIPOGTIOIOSNSSFAENPG LCGAPLKKKCGEEEDKEKKDKGLSWVAAGVYVPELFGALIGHILTSYKRWME WRFSCFSPLKK" /number=1 /gene="d13000w" 1753..3194 /number=1 3195..3260 /gene="d13000w"		
source	1..63289 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM020K8" 62805..140141 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM06G17" 134184..200576 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM08H22" 1753..4428 /gene="d13000w" join(1753..3194,3261..4428) /gene="d13000w" /note="similarity to Cf-2.2, Solanum pimpinellifolium" /codon_start=1 /product="disease resistance Cf-2 like protein" /protein_id="CAB10171.1" /db_xref="GI:2244748" /db_xref="SPTREMBL:O23253" /translation="MTITWISLCLIFCNSNLIYIAKDLCPDORDALFEKNESIP SPDSDLMLIQTAKARNNTDCCSWGIGSDPRTGVVVELDGNLGRNSSLF RLQHLQSIDLSTNDLSTLPDSSGNFKYRLVNLGLNFGELPTSLNLSLYLDDL SYNDLITGEIYDLSNGNLKRLRYLSLFCSEKPKIPSSICNLITLIDDLISMYTFEEL PDSKGNLSIRLVNLRHCRNFEGKIPSTGSLNLDLISKNEFTSEGDSSSLRL TDFQMLNLNLVNDLSSNQFKALPSPNMSLSLLEAFDISGNSFCSTIPSSLEML PSLIRKDLGTNDPSPGLKIGNISPSNLOELYIGENNINGPIPRSTLKLVGSAISL FMDTGIVDFISFQILKSLRSLDLSGINLISSSHLPSMMHLILSCNISOFPKEL ENQTSILYHLDISANOIEGOVPEMLRPLTSLFASDNKPSGEIPRAVCEIGTLVLSNN NFGSIPPCFEISNKTSLTSLIRNNSLSGVIPESLHGLIRSLIDVGSNLSQFPKSL INCSTIQLPLANTENKRIINDTFPSMLKSLPMLQVLVLRKSNFHGPITPSGSLSPKRL FDISNRSGLVPSDYFVGWVSWSLVDIIDMTPGFTVVGDOESFKSVLTIKGLN MELVSGEPIYKTIIVSGNRLBGDIPESIGIKELIVLMSNNAFGH1PSPISNLSN IQSLDLSQNRSLSGIPGELGELFTLARMFYSNMEGPIPOGTIOIOSNSSFAENPG LCGAPLKKKCGEEEDKEKKDKGLSWVAAGVYVPELFGALIGHILTSYKRWME WRFSCFSPLKK" /number=1 /gene="d13000w" 1753..3194 /number=1 3195..3260 /gene="d13000w"		
source	1..63289 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM020K8" 62805..140141 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM06G17" 134184..200576 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM08H22" 1753..4428 /gene="d13000w" join(1753..3194,3261..4428) /gene="d13000w" /note="similarity to Cf-2.2, Solanum pimpinellifolium" /codon_start=1 /product="disease resistance Cf-2 like protein" /protein_id="CAB10171.1" /db_xref="GI:2244748" /db_xref="SPTREMBL:O23253" /translation="MTITWISLCLIFCNSNLIYIAKDLCPDORDALFEKNESIP SPDSDLMLIQTAKARNNTDCCSWGIGSDPRTGVVVELDGNLGRNSSLF RLQHLQSIDLSTNDLSTLPDSSGNFKYRLVNLGLNFGELPTSLNLSLYLDDL SYNDLITGEIYDLSNGNLKRLRYLSLFCSEKPKIPSSICNLITLIDDLISMYTFEEL PDSKGNLSIRLVNLRHCRNFEGKIPSTGSLNLDLISKNEFTSEGDSSSLRL TDFQMLNLNLVNDLSSNQFKALPSPNMSLSLLEAFDISGNSFCSTIPSSLEML PSLIRKDLGTNDPSPGLKIGNISPSNLOELYIGENNINGPIPRSTLKLVGSAISL FMDTGIVDFISFQILKSLRSLDLSGINLISSSHLPSMMHLILSCNISOFPKEL ENQTSILYHLDISANOIEGOVPEMLRPLTSLFASDNKPSGEIPRAVCEIGTLVLSNN NFGSIPPCFEISNKTSLTSLIRNNSLSGVIPESLHGLIRSLIDVGSNLSQFPKSL INCSTIQLPLANTENKRIINDTFPSMLKSLPMLQVLVLRKSNFHGPITPSGSLSPKRL FDISNRSGLVPSDYFVGWVSWSLVDIIDMTPGFTVVGDOESFKSVLTIKGLN MELVSGEPIYKTIIVSGNRLBGDIPESIGIKELIVLMSNNAFGH1PSPISNLSN IQSLDLSQNRSLSGIPGELGELFTLARMFYSNMEGPIPOGTIOIOSNSSFAENPG LCGAPLKKKCGEEEDKEKKDKGLSWVAAGVYVPELFGALIGHILTSYKRWME WRFSCFSPLKK" /number=1 /gene="d13000w" 1753..3194 /number=1 3195..3260 /gene="d13000w"		
source	1..63289 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM020K8" 62805..140141 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM06G17" 134184..200576 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM08H22" 1753..4428 /gene="d13000w" join(1753..3194,3261..4428) /gene="d13000w" /note="similarity to Cf-2.2, Solanum pimpinellifolium" /codon_start=1 /product="disease resistance Cf-2 like protein" /protein_id="CAB10171.1" /db_xref="GI:2244748" /db_xref="SPTREMBL:O23253" /translation="MTITWISLCLIFCNSNLIYIAKDLCPDORDALFEKNESIP SPDSDLMLIQTAKARNNTDCCSWGIGSDPRTGVVVELDGNLGRNSSLF RLQHLQSIDLSTNDLSTLPDSSGNFKYRLVNLGLNFGELPTSLNLSLYLDDL SYNDLITGEIYDLSNGNLKRLRYLSLFCSEKPKIPSSICNLITLIDDLISMYTFEEL PDSKGNLSIRLVNLRHCRNFEGKIPSTGSLNLDLISKNEFTSEGDSSSLRL TDFQMLNLNLVNDLSSNQFKALPSPNMSLSLLEAFDISGNSFCSTIPSSLEML PSLIRKDLGTNDPSPGLKIGNISPSNLOELYIGENNINGPIPRSTLKLVGSAISL FMDTGIVDFISFQILKSLRSLDLSGINLISSSHLPSMMHLILSCNISOFPKEL ENQTSILYHLDISANOIEGOVPEMLRPLTSLFASDNKPSGEIPRAVCEIGTLVLSNN NFGSIPPCFEISNKTSLTSLIRNNSLSGVIPESLHGLIRSLIDVGSNLSQFPKSL INCSTIQLPLANTENKRIINDTFPSMLKSLPMLQVLVLRKSNFHGPITPSGSLSPKRL FDISNRSGLVPSDYFVGWVSWSLVDIIDMTPGFTVVGDOESFKSVLTIKGLN MELVSGEPIYKTIIVSGNRLBGDIPESIGIKELIVLMSNNAFGH1PSPISNLSN IQSLDLSQNRSLSGIPGELGELFTLARMFYSNMEGPIPOGTIOIOSNSSFAENPG LCGAPLKKKCGEEEDKEKKDKGLSWVAAGVYVPELFGALIGHILTSYKRWME WRFSCFSPLKK" /number=1 /gene="d13000w" 1753..3194 /number=1 3195..3260 /gene="d13000w"		
source	1..63289 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM020K8" 62805..140141 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM06G17" 134184..200576 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM08H22" 1753..4428 /gene="d13000w" join(1753..3194,3261..4428) /gene="d13000w" /note="similarity to Cf-2.2, Solanum pimpinellifolium" /codon_start=1 /product="disease resistance Cf-2 like protein" /protein_id="CAB10171.1" /db_xref="GI:2244748" /db_xref="SPTREMBL:O23253" /translation="MTITWISLCLIFCNSNLIYIAKDLCPDORDALFEKNESIP SPDSDLMLIQTAKARNNTDCCSWGIGSDPRTGVVVELDGNLGRNSSLF RLQHLQSIDLSTNDLSTLPDSSGNFKYRLVNLGLNFGELPTSLNLSLYLDDL SYNDLITGEIYDLSNGNLKRLRYLSLFCSEKPKIPSSICNLITLIDDLISMYTFEEL PDSKGNLSIRLVNLRHCRNFEGKIPSTGSLNLDLISKNEFTSEGDSSSLRL TDFQMLNLNLVNDLSSNQFKALPSPNMSLSLLEAFDISGNSFCSTIPSSLEML PSLIRKDLGTNDPSPGLKIGNISPSNLOELYIGENNINGPIPRSTLKLVGSAISL FMDTGIVDFISFQILKSLRSLDLSGINLISSSHLPSMMHLILSCNISOFPKEL ENQTSILYHLDISANOIEGOVPEMLRPLTSLFASDNKPSGEIPRAVCEIGTLVLSNN NFGSIPPCFEISNKTSLTSLIRNNSLSGVIPESLHGLIRSLIDVGSNLSQFPKSL INCSTIQLPLANTENKRIINDTFPSMLKSLPMLQVLVLRKSNFHGPITPSGSLSPKRL FDISNRSGLVPSDYFVGWVSWSLVDIIDMTPGFTVVGDOESFKSVLTIKGLN MELVSGEPIYKTIIVSGNRLBGDIPESIGIKELIVLMSNNAFGH1PSPISNLSN IQSLDLSQNRSLSGIPGELGELFTLARMFYSNMEGPIPOGTIOIOSNSSFAENPG LCGAPLKKKCGEEEDKEKKDKGLSWVAAGVYVPELFGALIGHILTSYKRWME WRFSCFSPLKK" /number=1 /gene="d13000w" 1753..3194 /number=1 3195..3260 /gene="d13000w"		
source	1..63289 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM020K8" 62805..140141 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM06G17" 134184..200576 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM08H22" 1753..4428 /gene="d13000w" join(1753..3194,3261..4428) /gene="d13000w" /note="similarity to Cf-2.2, Solanum pimpinellifolium" /codon_start=1 /product="disease resistance Cf-2 like protein" /protein_id="CAB10171.1" /db_xref="GI:2244748" /db_xref="SPTREMBL:O23253" /translation="MTITWISLCLIFCNSNLIYIAKDLCPDORDALFEKNESIP SPDSDLMLIQTAKARNNTDCCSWGIGSDPRTGVVVELDGNLGRNSSLF RLQHLQSIDLSTNDLSTLPDSSGNFKYRLVNLGLNFGELPTSLNLSLYLDDL SYNDLITGEIYDLSNGNLKRLRYLSLFCSEKPKIPSSICNLITLIDDLISMYTFEEL PDSKGNLSIRLVNLRHCRNFEGKIPSTGSLNLDLISKNEFTSEGDSSSLRL TDFQMLNLNLVNDLSSNQFKALPSPNMSLSLLEAFDISGNSFCSTIPSSLEML PSLIRKDLGTNDPSPGLKIGNISPSNLOELYIGENNINGPIPRSTLKLVGSAISL FMDTGIVDFISFQILKSLRSLDLSGINLISSSHLPSMMHLILSCNISOFPKEL ENQTSILYHLDISANOIEGOVPEMLRPLTSLFASDNKPSGEIPRAVCEIGTLVLSNN NFGSIPPCFEISNKTSLTSLIRNNSLSGVIPESLHGLIRSLIDVGSNLSQFPKSL INCSTIQLPLANTENKRIINDTFPSMLKSLPMLQVLVLRKSNFHGPITPSGSLSPKRL FDISNRSGLVPSDYFVGWVSWSLVDIIDMTPGFTVVGDOESFKSVLTIKGLN MELVSGEPIYKTIIVSGNRLBGDIPESIGIKELIVLMSNNAFGH1PSPISNLSN IQSLDLSQNRSLSGIPGELGELFTLARMFYSNMEGPIPOGTIOIOSNSSFAENPG LCGAPLKKKCGEEEDKEKKDKGLSWVAAGVYVPELFGALIGHILTSYKRWME WRFSCFSPLKK" /number=1 /gene="d13000w" 1753..3194 /number=1 3195..3260 /gene="d13000w"		
source	1..63289 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM020K8" 62805..140141 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM06G17" 134184..200576 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM08H22" 1753..4428 /gene="d13000w" join(1753..3194,3261..4428) /gene="d13000w" /note="similarity to Cf-2.2, Solanum pimpinellifolium" /codon_start=1 /product="disease resistance Cf-2 like protein" /protein_id="CAB10171.1" /db_xref="GI:2244748" /db_xref="SPTREMBL:O23253" /translation="MTITWISLCLIFCNSNLIYIAKDLCPDORDALFEKNESIP SPDSDLMLIQTAKARNNTDCCSWGIGSDPRTGVVVELDGNLGRNSSLF RLQHLQSIDLSTNDLSTLPDSSGNFKYRLVNLGLNFGELPTSLNLSLYLDDL SYNDLITGEIYDLSNGNLKRLRYLSLFCSEKPKIPSSICNLITLIDDLISMYTFEEL PDSKGNLSIRLVNLRHCRNFEGKIPSTGSLNLDLISKNEFTSEGDSSSLRL TDFQMLNLNLVNDLSSNQFKALPSPNMSLSLLEAFDISGNSFCSTIPSSLEML PSLIRKDLGTNDPSPGLKIGNISPSNLOELYIGENNINGPIPRSTLKLVGSAISL FMDTGIVDFISFQILKSLRSLDLSGINLISSSHLPSMMHLILSCNISOFPKEL ENQTSILYHLDISANOIEGOVPEMLRPLTSLFASDNKPSGEIPRAVCEIGTLVLSNN NFGSIPPCFEISNKTSLTSLIRNNSLSGVIPESLHGLIRSLIDVGSNLSQFPKSL INCSTIQLPLANTENKRIINDTFPSMLKSLPMLQVLVLRKSNFHGPITPSGSLSPKRL FDISNRSGLVPSDYFVGWVSWSLVDIIDMTPGFTVVGDOESFKSVLTIKGLN MELVSGEPIYKTIIVSGNRLBGDIPESIGIKELIVLMSNNAFGH1PSPISNLSN IQSLDLSQNRSLSGIPGELGELFTLARMFYSNMEGPIPOGTIOIOSNSSFAENPG LCGAPLKKKCGEEEDKEKKDKGLSWVAAGVYVPELFGALIGHILTSYKRWME WRFSCFSPLKK" /number=1 /gene="d13000w" 1753..3194 /number=1 3195..3260 /gene="d13000w"		
source	1..63289 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM020K8" 62805..140141 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM06G17" 134184..200576 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM08H22" 1753..4428 /gene="d13000w" join(1753..3194,3261..4428) /gene="d13000w" /note="similarity to Cf-2.2, Solanum pimpinellifolium" /codon_start=1 /product="disease resistance Cf-2 like protein" /protein_id="CAB10171.1" /db_xref="GI:2244748" /db_xref="SPTREMBL:O23253" /translation="MTITWISLCLIFCNSNLIYIAKDLCPDORDALFEKNESIP SPDSDLMLIQTAKARNNTDCCSWGIGSDPRTGVVVELDGNLGRNSSLF RLQHLQSIDLSTNDLSTLPDSSGNFKYRLVNLGLNFGELPTSLNLSLYLDDL SYNDLITGEIYDLSNGNLKRLRYLSLFCSEKPKIPSSICNLITLIDDLISMYTFEEL PDSKGNLSIRLVNLRHCRNFEGKIPSTGSLNLDLISKNEFTSEGDSSSLRL TDFQMLNLNLVNDLSSNQFKALPSPNMSLSLLEAFDISGNSFCSTIPSSLEML PSLIRKDLGTNDPSPGLKIGNISPSNLOELYIGENNINGPIPRSTLKLVGSAISL FMDTGIVDFISFQILKSLRSLDLSGINLISSSHLPSMMHLILSCNISOFPKEL ENQTSILYHLDISANOIEGOVPEMLRPLTSLFASDNKPSGEIPRAVCEIGTLVLSNN NFGSIPPCFEISNKTSLTSLIRNNSLSGVIPESLHGLIRSLIDVGSNLSQFPKSL INCSTIQLPLANTENKRIINDTFPSMLKSLPMLQVLVLRKSNFHGPITPSGSLSPKRL FDISNRSGLVPSDYFVGWVSWSLVDIIDMTPGFTVVGDOESFKSVLTIKGLN MELVSGEPIYKTIIVSGNRLBGDIPESIGIKELIVLMSNNAFGH1PSPISNLSN IQSLDLSQNRSLSGIPGELGELFTLARMFYSNMEGPIPOGTIOIOSNSSFAENPG LCGAPLKKKCGEEEDKEKKDKGLSWVAAGVYVPELFGALIGHILTSYKRWME WRFSCFSPLKK" /number=1 /gene="d13000w" 1753..3194 /number=1 3195..3260 /gene="d13000w"		
source	1..63289 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM020K8" 62805..140141 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM06G17" 134184..200576 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM08H22" 1753..4428 /gene="d13000w" join(1753..3194,3261..4428) /gene="d13000w" /note="similarity to Cf-2.2, Solanum pimpinellifolium" /codon_start=1 /product="disease resistance Cf-2 like protein" /protein_id="CAB10171.1" /db_xref="GI:2244748" /db_xref="SPTREMBL:O23253" /translation="MTITWISLCLIFCNSNLIYIAKDLCPDORDALFEKNESIP SPDSDLMLIQTAKARNNTDCCSWGIGSDPRTGVVVELDGNLGRNSSLF RLQHLQSIDLSTNDLSTLPDSSGNFKYRLVNLGLNFGELPTSLNLSLYLDDL SYNDLITGEIYDLSNGNLKRLRYLSLFCSEKPKIPSSICNLITLIDDLISMYTFEEL PDSKGNLSIRLVNLRHCRNFEGKIPSTGSLNLDLISKNEFTSEGDSSSLRL TDFQMLNLNLVNDLSSNQFKALPSPNMSLSLLEAFDISGNSFCSTIPSSLEML PSLIRKDLGTNDPSPGLKIGNISPSNLOELYIGENNINGPIPRSTLKLVGSAISL FMDTGIVDFISFQILKSLRSLDLSGINLISSSHLPSMMHLILSCNISOFPKEL ENQTSILYHLDISANOIEGOVPEMLRPLTSLFASDNKPSGEIPRAVCEIGTLVLSNN NFGSIPPCFEISNKTSLTSLIRNNSLSGVIPESLHGLIRSLIDVGSNLSQFPKSL INCSTIQLPLANTENKRIINDTFPSMLKSLPMLQVLVLRKSNFHGPITPSGSLSPKRL FDISNRSGLVPSDYFVGWVSWSLVDIIDMTPGFTVVGDOESFKSVLTIKGLN MELVSGEPIYKTIIVSGNRLBGDIPESIGIKELIVLMSNNAFGH1PSPISNLSN IQSLDLSQNRSLSGIPGELGELFTLARMFYSNMEGPIPOGTIOIOSNSSFAENPG LCGAPLKKKCGEEEDKEKKDKGLSWVAAGVYVPELFGALIGHILTSYKRWME WRFSCFSPLKK" /number=1 /gene="d13000w" 1753..3194 /number=1 3195..3260 /gene="d13000w"		
source	1..63289 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC TAM020K8" 62805..140141 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone		

Dx		MPI:	2002-217119/27.	
DR	P-FSDB;	AAM50937.		
Xx		Regulating gene expression in plants for controlling gene silencing,		
Pr		comprises altering the transcription or translation of an endonuclease		
Pt	nucleic acid sequence encoding a polypeptide comprising an exonuclease domain -			
Pt				
Ss	Claim 4; Page 93-94; 102pp; English.			
Xx		The present nucleotide sequence comprises cDNA encoding an		
Cc	Arbidopsis thaliana 3'-5' exonuclease domain (see AAM50937).			
Cc	specifically an RNase D related domain. The cDNA was identified			
Cc	on the basis of homology to the 3'-5' exonuclease sequence given in			
Cc	ABAY1795 via 3' and 5' RACE and TA-cloning. The sequence is nearly			
Cc	identical to an Arabidopsis exonuclease termed wrnexo. Nucleotide			
Cc	(see ABAY1795-807) and polyprotein (see AAM0927-39) sequences for			
Cc	3'-5' exonuclease domains were identified using different screening			
Cc	methods and search algorithms. The invention encompasses the			
Cc	suppression or increase of gene silencing in plants. This is			
Cc	achieved by altering the expression in the plant cell of a			
Cc	nucleotide sequence encoding a polypeptide comprising an exonuclease			
Cc	domain, preferably a 3'-5' exonuclease domain, and especially an			
Cc	RNase D-related domain. Expression of the nucleotide sequence is			
Cc	altered by altering its transcription or translation. It is			
Cc	reduced e.g. by sense suppression, antisense suppression, homologous			
Cc	recombination, the use of ribozymes, dominant-negative mutants,			
Cc	aptamers, zinc finger proteins, double-stranded RNA, or insertional,			
Cc	pilot or deletion mutagenesis, and increased by overexpression.			
Cc	The methods are useful for altering or stabilising the expression of			
Cc	a nucleotide sequence of interest in a plant cell or plant.			
Cc	Suppressing or decreasing expression of the nucleic acid molecule			
Cc	results in decreased levels of post transcriptional gene silencing			
Cc	and improved expression of genes of interest.			
SQ	Sequence 1049 BP; 292 A; 216 C; 220 G; 321 T; 0 other;			
	Query Match 100.0%; Score 1049; DB 24; Length 1049;			
	Best Local Similarity 100.0% ; Pred. No. 2.8e-297;			
	Matches 1049; Conservative 0; Mismatches 0; Indels 0; Gaps 0			
Oy	1 ACCAAGCATTAATTTTATTGTTTTGTTCAGTAAAGAATAATGCATCGTCAAAATTGA 60 			
Db	1 ACCTAACGATTAATTTTATTGTTTTGTTCAGTAAAGAATAATGCATCGTCAAAATTGA 60 			
Oy	TGACAGAGCCTTTTACAGAGAGAAGACTTCCTCCATFAGAGACCATFCGAAGCTTGCTACA 120 			
Db	TGACAGAGCCTTTTACAGAGAGAAGAGCTTCCTGGCATFAGAGGCCATFCGAAGCTTGCTACA 120 			
Oy	121 ATTTCCTCCCCTTCCTTCCTTCCTTCCTGCTGCTCCGACGCTACAAAGCTACAACT 180 			
Db	121 ATTTCTCCCCCTTCCTTCCTTCCTTCCTGCTGCTGCTCCGACGCTACAAAGCTACAACT 180 			
Oy	181 CGGTCCATGGCCAGAGAGAGATGCCAAATCCAATTCATATATCCGTGCCCAATTGC 240 			
Db	181 CCCTCCATGGCCAGAGAGAGATGCCAAATCCAATTCATATATCCGTGCCCAATTGC 240 			
Oy	241 CTCGTTCCATCACTCTTCCTCACATCTTTAAAAGATTTCTCTCCGTTCCGAGCTA 300 			
Db	241 CTCGTTCCATCACTCTTCCTCACATCTTTAAAAGATTTCTCTCCGTTCCGAGCTA 300 			
Oy	301 GGAAATTTTCCAGCATGAGAGTTGGTGTAGAGATTTGTATAGCAAAGACTGCTACTGAGG 360 			
Db	301 GGAAATTTTCCAGCATGAGAGTTGGTGTAGAGATTTGTATAGCAAAGACTGCTACTGAGG 360 			
Oy	361 TTGATTAACCGAGACAATGCAGCTTTATTAAGTTCTTGATACCAAGAGATGATCTGGAA 420 			
Db	361 TTGATTAACCGAGACAATGCAGCTTTATTAAGTTCTTGATACCAAGAGATGATCTGGAA 420 			
Oy	421 TAGCTTTTGTGGCTTGATATATGAGTGGAGACCAAGTTTATGAAGAAGGTGTTTCCCGG 480 			
Db	421 TAGCTTTTGTGGCTTGATATATGAGTGGAGACCAAGTTTATGAAGAAGGTGTTTCCCGG 480 			

OY		481	GGAAGGTGGCAGCGTCCACGATGATGTGAATAGTAATAATTATTTGAGATTTGCATATTT	540
Dd		481	GGAAAGTGTCACACTGTCTCACGATATGTAGTAGTAATATTTATTTGATTTGCATATTT	540
OY		541	TTCAATTCGTATCCCTCAAAGTCCAACAATCTTATTTGAAGATTCAACACTTTGAAAAG	600
Dd		541	TTCATTCGTATGCCCTCAAAGTCCCMAACATCTTATTTGAAGATTCAACACTTTGAAAAG	600
OY		601	TAGTATTGGAAATGATNGGTGACCTCTGTGAAGCTTTTTCCAGACTATGAGATTGATNCA	660
Dd		601	TAGTATTGGAAATGATNGGTGACCTCTGTGAAGCTTTTTCCAGACTATGAGATTGATNCA	660
OY		661	AAGAATGTGAGAGATCTTTTACAGATTTAGCCAACAAAAAATGTTGAGCATPAAAAATGGG	720
Dd		661	AAAGATGTTGAGAGATCTTTTACAGATTTAGCCAACAAAAAATGTTGAGCATPAAAAATGGG	720
OY		721	GCCTTGCGCTCACTAACTGAGACACTGTGTTGCCAAAGAGCCTCTGAAGCCAAACAGAAATCA	780
Dd		721	GCCTTGCGCTCACTAACTGAGACACTGTGTTGCCAAAGAGCCTCTGAAGCCAAACAGAAATCA	780
OY		781	GGCCTTGGAACCTGGAGATTATATCTCTGTGTCAAAGCAGCACTTCAATATGCGACAAACGG	840
Dd		781	GGCCTTGGAACCTGGAGATTATATCTCTGTGTCAAAGCAGCACTTCAATATGCGACAAACGG	840
OY		841	ATGCTTATGCTTCAATGAGCATCTTTTACAAGTCTTCTTAAGSACCCTTCGTATGCTGTACGTG	900
Dd		841	ATGCTTATGCTTCAATGAGCATCTTTTACAAGTCTTCTTAAGSACCCTTCGTATGCTGTACGTG	900
OY		901	GCTCATTAACGTGAAGAGGAAGCTTMAAGGTTAGCCTTTAAACCCCAAGAGTTAGCATCA	960
Dd		901	GCTCATTAACGTGAAGAGGAAGCTTMAAGGTTAGCCTTTAAACCCCAAGAGTTAGCATCA	960
OY		961	ATGATATGATACACCTATCTAGTCAAGTAGATGCAANTCTTGGAATATGTATCTAGT	1020
Dd		961	ATGATATGATACACCTATCTAGTCAAGTAGATGCAANTCTTGGAATATGTATCTAGT	1020
OY		1021	TCTGTGCTCCTTTAAACCGTCAGAAACTTAG	1049
Dd		1021	TCTGTGCTCCTTTAAACCGTCAGAAACTTAG	1049
			RESULT 2	
ID		AAC51089		
ID		AAC51089	standard; DNA; 1114 BP.	
XX		AAC51089;		
AC		18-OCT-2000	(first entry)	
DT				
XX				
DE		Arabidopsis thaliana	DNA fragment SEQ ID NO: 67235.	
XX				
KW		Hybridisation assay; genetic mapping; gene expression control;		
KW		protein identification; signal transduction pathway;		
KW		metabolic pathway; promoter; termination sequence; ss.		
OS		Arabidopsis thaliana.		
XX		EPI033405-A2.		
PV				
PD		06-SEP-2000.		
XX				
PF		25-FEB-2000; 2000EP-0301439.		
XX				
PR		25-FEB-1999; 99US-0121825.		
PR		05-MAR-1999; 99US-0123180.		
PR		09-MAR-1999; 99US-0123548.		
PR		23-MAR-1999; 99US-0125788.		
PR		25-MAR-1999; 99US-0126264.		
PR		29-MAR-1999; 99US-0126785.		
PR		01-APR-1999; 99US-0127462.		
PR		06-APR-1999; 99US-0128234.		
PR		08-APR-1999; 99US-0128714.		


```
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160776.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161982.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

Query Match 99.0%; Score 1039; DB 21; Length 1114;
Best Local Similarity 100.0%; Pred. No. 2.4e-294;
Matches 1039; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 11 TAAATTTTATTTTGTCTTCAAGTAAAGAAATGTCATGCTCAATTTGATCGACGAGC 70
DB 1 TAAATTTTATTTTGTCTTCAAGTAAAGAAATGTCATGCTCAATTTGATCGACGAGC 60
QY 71 TTTTACAGAGAGAGAGCTTCGCTACGACGACGAGAGAGCTTCACATTTCCCG 130
DB 61 TTTTACAGAGAGAGAGCTTCGCTACGACGACGAGAGAGCTTCACATTTCCCG 120
QY 131 TTCTTCTTCTTCTTCTCTCTGCTCGACGAGTACCACTACCACTCCGTCATG 190
DB 121 TTCTTCTTCTTCTTCTCTCTGCTCGACGAGTACCACTACCACTCCGTCATG 180
QY 191 CCACGAGAGAGATCCAAATTAATCCCAATTAATCCGCGCAATTCCTCGTTCAT 250
DB 181 CCACGAGAGAGATCCAAATTAATCCCAATTAATCCGCGCAATTCCTCGTTCAT 240
QY 251 CACTTCTTCACTTATTAAGATTTCTCTCCCGTTGCCGAGCTAGAAATTTCC 310
DB 241 CACTTCTTCACTTATTAAGATTTCTCTCCCGTTGCCGAGCTAGAAATTTCC 300
QY 311 AGCAATGAGGTTGGTGTAGATTTGTATAGCAAGACTGCTAGAGTTGATAAGC 370
DB 301 AGCAATGAGGTTGGTGTAGATTTGTATAGCAAGACTGCTAGAGTTGATAAGC 360
QY 371 AGCAATGAGGTTGGTGTAGATTTGTATAGCAAGAGATGATCTGATTTGAT 430
DB 361 AGCAATGAGGTTGGTGTAGATTTGTATAGCAAGAGATGATCTGATTTGAT 420
QY 431 TGGCTTGATATTTAGTGTAGAGACCAAGTTTAAAGAGTCTCCCGGGAGAGTTGC 490
DB 421 TGGCTTGATATTTAGTGTAGAGACCAAGTTTAAAGAGTCTCCCGGGAGAGTTGC 480
QY 491 GACTGTCCAGATATGTGTAGATATATTTGTATGATTTATGCATATTTTCAATCTGC 550
DB 481 GACTGTCCAGATATGTGTAGATATATTTGTATGATTTATGCATATTTTCAATCTGC 540
QY 551 TATCCCTCAAGCTCTCCACATCTTATGTGAAGTTCAACCTTTGAAGTAGTATGG 610
DB 541 TATCCCTCAAGCTCTCCACATCTTATGTGAAGTTCAACCTTTGAAGTAGTATGG 600
QY 611 AATTGATGTGATCTGTGAAGCTTTTCCATGATGTGATGATAGTATCAAGATGTTGA 670
DB 601 AATTGATGTGATCTGTGAAGCTTTTCCATGATGTGATGATAGTATCAAGATGTTGA 660
```

```
QY 671 GGATCTTTAGATTTTACCAACCAAAAAATTTGGTGTAGATAAAAATGGGCGCTTGCCTC 730
DB 661 GGATCTTTAGATTTTACCAACCAAAAAATTTGGTGTAGATAAAAATGGGCGCTTGCCTC 720
QY 731 ACTTACTGAGACACTTTTGGCAAGAGCTCTGAGACCCAAACAGATCAGCTTGGGAA 790
DB 721 ACTTACTGAGACACTTTTGGCAAGAGCTCTGAGACCCAAACAGATCAGCTTGGGAA 780
QY 791 CTGGAGATTTTATTCCTGTCAAAAGCAGCAGTTTACATFAGCAGCAAGGATGCTTATGC 850
DB 781 CTGGAGATTTTATTCCTGTCAAAAGCAGCAGTTTACATFAGCAGCAAGGATGCTTATGC 840
QY 851 TTCTAGCATCTTTTACAAAGGTTTCTTAAAGACCTTCTGATGCTCAGTGCATTAACG 910
DB 841 TTCTAGCATCTTTTACAAAGGTTTCTTAAAGACCTTCTGATGCTCAGTGCATTAACG 900
QY 911 TGAAGAGAGAGCTTAAAGGTTAGCCTATATTAACCCAAAGATTTAGCATCAATGATATGAT 970
DB 901 TGAAGAGAGAGCTTAAAGGTTAGCCTATATTAACCCAAAGATTTAGCATCAATGATATGAT 960
QY 971 ACACCTAATCTAGTCAAGTAGATGATCAATTTCTGTGAATATTTGATCTAGTTCGTCCT 1030
DB 961 ACACCTAATCTAGTCAAGTAGATGATCAATTTCTGTGAATATTTGATCTAGTTCGTCCT 1020
QY 1031 TTACCGTCCAGAAACTAG 1049
DB 1021 TTACCGTCCAGAAACTAG 1039
```

```
RESULT 3
AAC40589
ID AAC40589 standard; DNA; 1129 BP.
XX
AC AAC40589;
XX
DE 17-OCT-2000 (first entry)
XX
KW Arabidopsis thaliana DNA fragment SEQ ID NO: 28855.
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
```

PR 07-MAY-1999; 99US-013863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136782.
PR 28-MAY-1999; 99US-0137222.
PR 01-JUN-1999; 99US-0137528.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145313.
PR 27-JUL-1999; 99US-0145318.
PR 27-JUL-1999; 99US-0145319.
PR 28-JUL-1999; 99US-0145351.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158822.
PR 12-OCT-1999; 99US-0158832.
PR 13-OCT-1999; 99US-0158969.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.

```

PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match 95.5%; Score 1002.2; DB 21; Length 1129;
Best Local Similarity 97.8%; Pred. No. 1.5e-283;
Matches 1016; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

```

```

QY 11 TAAATTTTATTTTGTCTTCTAGTAAAGAAATGTCATGCTCAATTTGATCGACGACG 70
DB 2 TTAATTTGTCTTTTCTTCTTCAATTAAGAAATGTCATGCTCAATTTGATCGACGACG 61
QY 71 TTTTACAGAGAGAGGCTTCGCTATCGACGCCATCGAGCTTCCATTTCTCCG 130
DB 62 TTTTACAGAGAGAGGCTTCGCTATCGACGCCATCGAGCTTCCATTTCTCCG 121
QY 131 TTTCTTCTTCTTCTTCTCTGCTGCTCGACGCTACAGCTACACACTCCGTCATGG 190
DB 122 TTTCTTCTTCTTCTTCTCTGCTGCTCGACGCTACAGCTACACACTCCGTCATGG 181
QY 191 CCACGAGAGAGATCCAAATCAATCCCAATTAATATCCGTCGCAATTTGCTGTCAT 250
DB 182 CCACGAGAGAGATCCAAATCAATCCCAATTAATATCCGTCGCAATTTGCTGTCAT 241
QY 251 CACTTCTTCTACATCTTAATAAGATTTCTCTCTCCGTTCCGACGCTAGAAATTTCC 310
DB 242 CACTTCTTCTACATCTTCTTAAGATTTCTCTCTCCGTTCCGACGCTAGAAATTTCC 301
QY 311 AGCAATGAGGTTGGTGGTAGATTTTGTATAGCAAGACTGCTACTGAGTTGATTAAGC 370
DB 302 AGCAATGAGGTTGGTGGTAGATTTTGTATAGCAAGACTGCTACTGAGTTGATTAAGC 361
QY 371 AGCAATGAGGTTGGTGGTAGATTTTGTATAGCAAGACTGCTACTGAGTTGATTAAGC 430
DB 362 AGCAATGAGGTTGGTGGTAGATTTTGTATAGCAAGACTGCTACTGAGTTGATTAAGC 421
QY 431 TGGCTTGGATTTGAGTGGAGACCAAGTTTAAAGAGTGTCTCCGCGGAAGGTTGC 490
DB 422 TGGCTTGGATTTGAGTGGAGACCAAGTTTAAAGAGTGTCTCCGCGGAAGGTTGC 481
QY 491 GACTGTCCAGATATGTGTAGATAGTAATTAATTTGTATGATTAATTTTCAATTCG 550
DB 482 AACTGTCCAGATATGTGTAGATAGTAATTAATTTGTATGATTAATTTTCAATTCG 541
QY 551 TATTCCTCAAGTCTCCACATCTTATTAAGATTCAACCTTGTAAAGTAGTATGG 610
DB 542 TATTCCTCAAGTCTCCACATCTTATTAAGATTCAACCTTGTAAAGTAGTATGG 601
QY 611 AATTGATGTGACTCTGTGAAGCTTTTCCATGACTATGAGATTAGTCAAAAGTGTGA 670
DB 602 AATTGATGTGACTCTGTGAAGCTTTTCCATGACTATGAGATTAGTCAAAAGTGTGA 661
QY 671 GGATCTTTTCAGATTTAGCCCAACCAAAATTTGGTGAATATAAATTTGGGCTTGCCTC 730
DB 662 GGATCTTTTCAGATTTAGCCCAACCAAAATTTGGTGAATATAAATTTGGGCTTGCCTC 721
QY 731 ACTTACGAGACACTTGTTCGAAAGAGCTCTGAGACCAAAAGAAATCAGGCTTGGGAA 790
DB 722 ACTTACGAGACACTTGTTCGAAAGAGCTCTGAGACCAAAAGAAATCAGGCTTGGGAA 781
QY 791 CTGGAGATTATTCCTGTCAAGAGCAGATTACATAGCAGCAAGGATGCTTATGC 850
DB 782 CTGGAGATTATTCCTGTCAAGAGCAGATTACATAGCAGCAAGGATGCTTATGC 841

```

```

QY 851 TTCATGCACTTTTACAGGTTCTTAAGACCTTCTGATGCTGATGCTCATTAACG 910
DB 842 TTCATGCACTTTTACAGGTTCTTAAGACCTTCTGATGCTGATGCTCATTAACG 901
QY 911 TGAAGAGAGAGCTTAAAGTTAGCCCTATTAACCCCAAGAGTTAGCATCAATGATGAT 970
DB 902 TGAAGAGAGAGCTTAAAGTTAGCCCTATTAACCCCAAGAGTTAGCATCAATGATGAT 961
QY 971 ACACCTATCTAGTCAAGTATGATGCAATTTCTTGTGATATGATCTAGTTCGTCCT 1030
DB 962 ACACCTATCTAGTCAAGTATGATGCAATTTCTTGTGATATGATCTAGTTCGTCCT 1021
QY 1031 TTAACGCTCCAGAACTAG 1049
DB 1022 TTAATGCTCCAGAACTAG 1040

RESULT 4
ABA91795
ID ABA91795 standard; cDNA; 942 BP.
XX
AC ABA91795;
XX
XX 15-MAY-2002 (first entry)
DE Arabidopsis RNase D related domain (3'-5' exonuclease domain) DNA.
DE RNase D; enzyme; 3'-5' exonuclease; gene expression; plant;
KW gene silencing; transgenic plant; gene; ss.
XX Arabidopsis thaliana.
XX
FH Key Location/Qualifiers
FT CDS 1..942
FT /tag= a
FT /product= "3'-5' exonuclease domain"
XX
PN W020210362-A2.
XX
PD 07-FEB-2002.
XX
PF 30-JUL-2001; 2001WO-EP08825.
XX
PR 01-AUG-2000; 2000US-222202P.
XX
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
PA (NOVS ) NOVARTIS FORSCHUNGSSTIFTUNG ZWEIGNIEDERL.
XX
PI Levin JZ, Phillips KL, Budziszewski GJ, Meins F, Glazov EA.
XX
DR WPI: 2002-21719/27.
XX
DR P-PSDB: AAM50927.
XX
PT Regulating gene expression in plants for controlling gene silencing,
PT comprises altering the transcription or translation of an endonuclease
PT nucleotide sequence encoding a polypeptide comprising an exonuclease
PT domain
XX
XX
PS Claim 4; Page 70; 102pp; English.
XX
XX The present nucleotide sequence encodes an Arabidopsis thaliana 3'-5'
XX exonuclease domain (see AAM50927), specifically an RNase D related
XX domain (Genpept accession CAB36851). The nucleotide sequence is
XX found in BAC F18A5, GenBank accession number AL035528.2. Nucleotide
XX (see ABA91795-807) and polypeptide (see AAM50927-39) sequences for
XX 3'-5' exonuclease domains were identified using different screening
XX methods and search algorithms. The invention encompasses the
XX suppression or increase of gene silencing in plants. This is
XX achieved by altering the expression in the plant cell of a
XX nucleotide sequence encoding a polypeptide comprising an exonuclease
XX domain, preferably a 3'-5' exonuclease domain, and especially an
XX RNase D-related domain. Expression of the nucleotide sequence is

```


SQ Sequence 4299 BP; 1400 A; 759 C; 950 G; 1190 T; 0 other;

Query Match	10.6%	Pred. 111;	DB 24;	Length 4299;
Best Local Similarity	54.1%	Pred. No. 9.1e-22;		
Matches 272;	Conservative	0;	Mismatches 225;	Indels 6;
				Gaps 2

QY	396	GATACCAAGAGAGATGAATGAACTGGAAATACCTTTTGTGGCTGGAAATTTGGAGTGGACCA	455
Db	202	GAAATTTAGCATGAGCTATACAGATGGGAGATGGTGGGATTTTGACATGGAGTGGCCACCA	261
QY	456	AGTTTATGAAAAAGTGTCTCCCGGGGAAGGTGGCACTGTCAGATATATGTAGATAGT	515
Db	262	TTATACAAATAGAGGCAAACT--TGCCAAAGTGCACATAATTCAGTTGTGTTCGAG	318
QY	516	AATATTTGTGATGATTATGCATATTTT--CATCTGATGCCCAAGATCCCAACAT	572
Db	319	AGCAAAATTTACTTGTTCACGTTTCTTCACATGTCAGTATTTTCCCGAGGATTAATAATG	378
QY	573	CTTATTTGAAGATTCAACACTTGTAAAGTGTAGTATGTGAATTTGATGTGACTGTGAG	632
Db	379	TTGCTTGAAAAATAAAGCAGCTTAAAAAGCCAGGTGAGAAATTTGAAGGAGATCACTGGAA	438
QY	633	CTTTTCATGACATATGAGAGTTGTATCCAAAGATGTGAGATCTTTCAGATTTATGCCAAC	692
Db	439	CTTCTACGAGACTTTGATATTCAAATTTGAAGAATTTTGTGGAGTTGACAGATGTTGCCAAT	498
QY	693	CAAAAAATTTGTGTGAGATTAATAAATGGGGCCTTGGCTTCACACTGACACTGTTTGC	752
Db	499	AAAAACCTGAATGTACAGAGACTGTGAGACTTACACATGTGTGTTAAACACCTTTAGGT	558
QY	753	AAAGACTCTCTAAGCCAAACAGAATTCAGGCTTGGGACTGGGAGTTTATCCTCTGTCA	812
Db	559	AAACACTCTCTGAAGACAAAGTCTATCCGCTGTACCAATTTGAGATTAATTTCTCTACT	618
QY	813	AACGACATTTCAATATACACAGCAACGAGATGTTATGCTTCATGGCACTTTCAAGATT	872
Db	619	GAGACCAAAACTGTATGCACACCACTGATGCTTATGTGTTTATATTTACCGAAAT	678
QY	873	CTTAAGGACCTTCGATGCTGT	895
Db	679	TTAGAGATTTTGGATGATACGT	701

RESULT 6

ID AAV65701 standard; cDNA to mRNA; 5189 BP.

AC AAV65701;

DT 10-DEC-1998 (first entry)

DNA fragment of a Werner's syndrome gene.

KW Werner's syndrome; diagnosis; ds.

OS Homo sapiens.

PN JP10201498-A.

PD 04-AUG-1998.

PF 24-JAN-1997; 97JP-0011268.

PR 24-JAN-1997; 97JP-0011268.

PA (EIJ-) EIJIN KENKYUSHO KK.

DR WPI; 1998-474499/41.

PT	Detection of mutation in gene causing human Werner's syndrome - and
PT	oligo:nucleotide used for detection, comprises amplifying DNA and
PT	synthesising oligo:nucleotide
...	

PS Claim 1; Pages 10-11; 17pp; Japanese.

The present sequence represents a gene, the mutation of which causes Werner's syndrome. The specification describes the detection of a mutation in a gene causing human Werner's syndrome. The method comprises amplifying a DNA fragment containing a mutation at position 733, 734, 1630 or 1446 of AW65701 or at position 42 of AW65702 and synthesising an oligonucleotide so that the base at the above site comes to be the 3' end based on the base sequence of AW65701-02, or an oligonucleotide in which the base adjacent to the 3' end comes to be the 5' end. The oligonucleotides are hybridised with the resultant amplified fragment. The method can be used to diagnose Werner's syndrome.

SQ Sequence 5189 BP; 1663 A; 889 C; 1129 G; 1508 T; 0 other;

Query Match	10.6%	Score 111	DB 19	Length 5189
Best Local Similarity	54.1%	Pred. No. 9.9e-22		
Matches 272; Conservative	0	Mismatches 225	Indels 6	Gaps 2

QY	396	GATACCAGAGAGATGAATCTGGAAATAGCTTTTGTGGCTGGATATTTAGTGAACCA	455
Db	433	CATATTACATACAGCTCATACATGGGGAATGTGTGGATTTTGACATGGAGTGGACCA	492
QY	456	AGTTTAAAGAAAGTGTTCTCCCGGGGAAGTTGCGACTGTCACATATGTAGATAGT	515
Db	493	TTATCAATAGAGGGGAAACT---TGCCAAAGTTGCACTAATTCAGTTGTGTCTGTGAG	549
QY	516	AATTAATGTGATGATATGCAATATTTT---CATTCGGATATCCCTCAAGTCCCAACAT	572
Db	550	AGCAATGTACTTGTTCACGTTTCTTCCATGTCAAGTTTTCGCCAGGATTAATAATG	609
QY	573	CTTATTTGAAGATTCAACCTGTAAAGTAGGTATTGGAATTTGATGGTCACTCTGTGAAG	632
Db	610	TTGCTTTGAAATAAAGACAGTTAAAAAGCAGCGTGTAGCAATTTGAAGAAATCAGTGGAAA	669
QY	633	CTTTCCCAATGATATAGAGATTAGTATCAAGATGTGAGAGATCTTTCAGATTAGCAAC	692
Db	670	CTTCTACGGACTTTGGATATCAAAATTTGAAGAAATTTGTGTGAGTTGACAGATGTGCCAAT	729
QY	693	CAAAAAATTTGTTGGAGATTAATAAATATGGGGCTTGTGCTCACTAACTAGACACTTGTTCG	752
Db	720	AAAAAGCTGAATTTGATACAGAGACCTGGAGACCTTAACAGTCTGGTTAAACACCTCTAGGG	789
QY	753	AAAGAGCTCTCTGAAGCCAAACAGAATTCAGGCTTTGGAACTGGGAGTTTATCTCTGTCA	812
Db	790	AAACAGCTCTCGAAAAGACAAAGTCTATTCGCTGTAGCAATTTGGAGTAAATTTCTCTCACT	849
QY	813	AAGCAGAGTTTCAATATACGACCAACGAGATGCTTATAGCTTCAATGGCATTTTACAAGTT	872
Db	850	GAGGACCAAGAAACTGTATCCAGCACCACTGATGCTTATAGCTGTTTATTTATTTACCGAAAT	909
QY	873	CTTAAAGGACCTCTCGATGCTGT	895
Db	910	TTAGAGATTTTGGATGATACGT	932

RESULT 7

ID ABL65821 standard; DNA; 5189 BP.

AC ABL65821;

DT 15-MAY-2002 (first entry)

DE Lung cancer related gene sequence SEQ ID NO:4158.

KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; Lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma
 KW gene; ds.

OS Homo sapiens.

PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234505P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237178P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237605P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PA Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S,
PI Soppet DR, Weaver Z;
XX
DR WPI: 2002-188264/24.
XX
XX
PT Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set
XX
PS Claim 1; SEQ ID 4158; 44pp: English.
XX
XX The present invention describes a method (M1) for screening for an
XX anti-neoplastic agent. The method involves exposing cells to a chemical
XX agent to be tested for anti-neoplastic activity, determining a change in
XX expression of at least one gene (I) of a signature gene set, where (I)
XX comprises a sequence (S) selected from 8447 sequences (given in ABL1664
XX to ABL70110), or is at least 95% identical to (S), where a change in
XX expression is indicative of anti-neoplastic activity. (I) has cytosol
XX activity and can be used in gene therapy. M1 can be used for screening
XX an anti-neoplastic agent, and can be used for producing a product which
XX is the data collected with respect to the anti-neoplastic agent as a
XX result of M1, and the data is sufficient to convey the chemical

CC	structure and/or properties of the agent. M1 can be used in the
CC	treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC	oesophageal, ovarian, kidney, prostate or pancreatic cancer.
CC	adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC	infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC	carcinoma, papillary carcinoma and M1M's tumour.
XX	
SQ	Sequence 5189 BP; 1663 A; 889 C; 1129 G; 1508 T; 0 other;
Query Match	10.6%; Score 111; DB 24; Length 5189;
Best Local Similarity	54.1%; Pred. No. 9,9e-22;
Matches 272;	Conservative 0; Mismatches 225; Indels 6; Gaps 2
YY	
Db	
YQ	396 GATACCAAGAGAGATGATCTGGAAATACGTTTGTGGCTTGATATTGACTGGAGCCA 455
	433 GATATTAGCATGAGTCATCAGATGGGATGTGGTGATTTGCATGAGAGGCCACA 492
YQ	456 AGTTTTAGAAAAGTGTCTCCCGGGGAAAGTTGGAGATGTCCAGATATGTGATAGT 515
Db	493 TTATACATATAGAGGAAACT--TGGCAAAGTGCATAATTCAGTTGTGGTTCTGAG 549
YQ	516 AATATTGTGATGTATTCGATATTTT---CATTCGTGATCCCTCAAAGCTCCACAT 572
Db	550 AGCAATGTTACTGTCTGCACGTTCTTCOCATGTACGATTTTCCCAGGATTAATAATG 609
YQ	573 CTATTAGAAGATTCAACACTTGTAAAGTAGTATTTGGAATGTATGTGTGACTGTGAG 632
Db	610 TTGCTTGAAAAATTAACGACGTAAAAAAGGAGGTGAGAAATGGAAGGATCACTGAAA 669
YQ	633 CTTTTCATGACTATGAGAGTGTATCAAAAGATGTGAGGATCTTTACGATTAGCCAC 692
Db	670 CTTCCTCGAGCTTGTATTCAAATTTGAAGATTTTGTGGAGTTGACAGATGTGCCAAT 729
YQ	693 CAATAAATTTGNTGAGATTAATAATTTGGGGCTTGGCTCACATCTGAGACACTGTTGC 752
Db	730 AAAAACGTAAATGTACAGAGACCTTGGACCCTTAAACAGTCTGGTTAAACACCTTTAGGT 789
YQ	753 AAAGAGCTCCTGAAGCCAAACAGAAATCAGGCTTGGGAATGGAGTTTATCTCTGTCA 812
Db	790 AAACAGCTCCTGAAAGACAAAGTCTATCCGCTGTACAAATGTGAGTAATTTCTCTCAGT 849
YQ	813 AACGACAGTTCATATACGACCAACGATGCTTATGCTTATGSCATCTTTACAGGTT 872
Db	850 GAGAGCCCAAAACCTGTATCAGCCACTGATGCTTATGCTTATTTATTTACGGAAT 909
YQ	873 CTTAAGAGACTTCCGTGATGCTGT 895
Db	910 TTAGAGATTTTGGATGATCTGT 932
RESULT 8	
AAK83001	
ID	AAK83001 standard; DNA; 5208 BP.
XX	
AC	AAK83001:
XX	
DT	31-AUG-1999 (first entry)
XX	
DE	Human WRN coding sequence.
XX	
KW	Human; WRN; Werner's syndrome; detection; diagnosis; autosomal;
KM	recessive disorder; phenotype; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO9724435-A1.
XX	
PD	10-JUL-1997.
XX	
PF	30-DEC-1996; 96MO-US20785.
XX	
XX	12-APR-1996; 96US-0632175.
RR	29-DEC-1995; 95US-0009409.

[illegible]

XX	RESULT 10
XX	AAV35131
ID	AAV35131 standard; cDNA to mRNA; 1164 BP.
XX	
AC	AAV35131;
XX	
DT	28-AUG-1998 (first entry)
XX	
DE	Mouse WRN helicase RACE fragment cDNA #3.
XX	
KW	Werner's syndrome; WRN; helicase; murine; diagnosis; RACE; ds.
XX	
OS	Mus sp.
XX	
PN	JP10146188-A.
XX	
PD	02-JUN-1998.
XX	
PF	15-NOV-1996; 96JP-0304721.
XX	
PR	15-NOV-1996; 96JP-0304721.
XX	
PA	(EIJ1-) EIJIN KENKUSHO KK.
XX	
DR	WPI; 1998-369863/32.
XX	
PT	New mouse gene encoding protein - used for diagnosis of human
PT	Werner's syndrome
XX	
PS	Example 1; Page 26-27; 30pp; Japanese.
XX	
CC	This sequence is a fragment of novel mouse WRN helicase protein which
CC	is associated with Werner's syndrome and has been generated by RACE
CC	(rapid amplification of cDNA ends). The mouse WRN gene is useful for
CC	the diagnosis of human Werner's syndrome.
XX	
SO	Sequence 1164 BP; 343 A; 223 C; 288 G; 310 T; 0 other;

Query Match	7.7%	Score 81.2	DB 19,	Length 1164,
Best Local Similarity	52.9%	Pred. No. 2.7e-13;		
Matches 198; Conservative	0;	Mismatches 173;	Indels 3;	Gaps 1

QY	481	GGAGGTTCCGACTCCGAGTATGTAGAGTAAATATGATGATTATGCATTTT	540
Db	449	GCAGATCCGATATCCAGTTGTGTGTCTGAGAACAAATGTTACTGTTCACATT	508
QY	541	TT---CATTCGTGTAATCCCTCAAACTCCAAACATCTTATTGAAGATTCAACTGTAA	597
Db	509	CTTCCATGTCATTTTCCCTCCAGGATTAATAAATGTTACTGAAACAAATTCATTTAGA	568
QY	598	AGTAGGTTTGGATTGATGTAAGCTTTGGAAGCTTTTCATGACTATGAGATTAGTA	657

Db	569	AGGCAGGGCGTTGGGANTGAAGGGACCAAGTGGAAACTTCGCGTGATATTTGACGTCAAGT	628
OY	658	TCAAAGATGTTGAGGAGTCTTTACAGATTTAGCCACCAAAAAATTGGTGGAGATATAAAAT	717
Db	629	TGGAGAGTTTTTTGGAGCTGACGGAGATGTTGCCAAATGAATAATTGAATGCCACAGACT	688
OY	718	GGGGCCCTTGCCCTACACACTGAGACACTGTTTGCAGAGAGCTCCAGAGCCAACAGAA	777
Db	689	GGAGCCCAATGCTCTGCTTAAACAACGCTTTHGGGAAACACTTTTGAATAACAAGTCCA	748
OY	778	TCAAGCTTGGGAACCTGGAGATTTTATCTCTGTCAAAGCAGCAGATTACATATCCAGCAA	837
Db	749	TCCGCTGCACGCAATTTGGAGTAATTTCCCCCTCCTACGTAGAGCAGCAAACTGTATCGACCA	808
OY	838	CGAGTGCCTATGCT	851
Db	809	CTGATGCTATGCT	822

XX	RESULT 11
XX	AAV35115
ID	AAV35115 standard; cDNA to mRNA, 4206 BP.
XX	
AC	AAV35115;
XX	
DT	28-AUG-1998 (first entry)
XX	
DE	Mouse WRN helicase cDNA.
XX	
KW	Werner's syndrome; WRN; helicase; murine; diagnosis; ds.
XX	
CS	Mus sp.
XX	

EH	Key	Location/qualifiers
FT		1..4206
FT	CDS	/*tag= a
FT		/product= WRN helicase
FT		

PN	JP10146188-A.
XX	
PD	02-JUN-1998.
XX	
PF	15-NOV-1996; 96JP-0304721.
XX	
PR	15-NOV-1996; 96JP-0304721.
XX	
PA	(EIJ1-) EIJIN KENKYUSHO KK.
XX	
DR	WPI; 1998-369863/32.

PT New mouse gene encoding protein - used for diagnosis of human
PT Werner's syndrome
XX
PS Claim 2; Page 22-23; 30pp; Japanese.
XX

CC This sequence encodes a novel mouse WRN helicase protein which is
CC associated with Werner's syndrome. The mouse WRN gene is useful for
CC the diagnosis of human Werner's syndrome.

SQ Sequence 4206 BP; 1284 A; 876 C; 976 G; 1070 T; 0 other;

Query Match	7.7%	Score	81.2	DB	19	length	4206
Best Local Similarity	52.9%	Pred	No. 5e-13				
Matches 198, Conservative						Indels	3
						Gaps	1

Oy	48	GGAGGTGGGAATCGCCAGATATGTTGATAGTAATATATGGAGTTATGCAATTT	540
	266	GCAGATCCAGATGACCAAGTGTGTGTCTCGAACAACAAATGTTACTTGTTCACATTT	3255
Db	541	TT---CAATTCGTGATTCCTCAAAGTCTCCAAACATCTTATTTGAAGATTCACACTGTAA	597
Oy	326	CTTCGCATGTCAGTTTCCCCAGGGATTAAAAATGTTCTGTGAAGAACAAATTCATTATAGA	3855

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2003, 06:36:38 ; Search time 1820 Seconds
(without alignments)
9334.651 Million cell updates/sec

Title: US-09-896-186b-23
Perfect score: 1049
Sequence: 1 accaagcattatattttat.....tttaacgcgtcagaactag 1049

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST: *
1: em_estba:*
2: em_esthum:*
3: em_estln:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_pod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	288.8	27.5	435	9	AU226180 AU226180
2	273.8	26.1	840	10	BE659005 BE659005
3	261	24.9	704	12	BG441595 GA_Ea001
4	245	23.4	552	10	AM201789 AM201789
5	243.8	23.2	671	9	AU237507 AU237507
6	241.2	23.0	467	10	AM832139 AM832139

7	235.4	22.4	661	14	BO799068	EST 1237
8	233.8	22.3	536	13	BM406124	EST580451
9	223	21.3	505	12	BG133320	EST466112
10	213.6	20.4	531	13	BM094501	EST156703
11	172.2	16.4	515	13	BM358892	GA_Ea001
12	169.6	16.2	517	13	BM358891	GA_Ea001
13	167.4	16.0	555	14	BM731125	EST168a05
14	162	15.4	595	13	Bj213300	Bj213300
15	156.8	14.9	499	13	BI433038	EST535799
16	136.6	14.9	536	14	BM890816	EST08c02
17	156	14.9	563	13	BM527093	sal149e10
18	152	14.5	531	10	AM460204	sh56c10.y
19	151	14.4	684	12	BG589996	EST497838
20	143.2	13.7	625	12	BE999775	EST431498
21	142.4	13.6	560	12	BG157629	sa88c03
22	140.8	13.4	572	9	AU235472	AU235472
23	134.4	12.8	680	13	Bj321563	Bj321563
24	129.6	12.4	407	12	BE821184	GM700024A
25	129	12.3	220	17	BH849559	SAUK 0698
26	124.6	11.9	431	9	AT748378	sb52a06.y
27	117	11.2	445	10	AM119478	sd46a06.y
28	100.6	9.6	494	10	AM720170	LJNEST16b
29	99	9.4	654	13	Bj220795	Bj220795
30	97.8	9.3	660	17	B61114	T20K8TF TAM
31	96	9.2	600	13	Bj316072	Bj316072
32	84.6	8.1	550	13	BM785803	K-EST0064
33	76.2	7.3	405	17	CNS00V8H	AT091967
34	68.8	6.4	533	17	AQ580338	AT091967
35	65.4	6.2	473	17	AQ866027	nbe00026P
36	65.4	6.2	901	17	AQ867999	nbe00018H
37	62	5.9	342	17	AO841745	TI32158b
38	62	5.9	457	17	BH168898	SAUK 0003
39	57.2	5.5	447	13	BI420478	LJNEST57f
40	56.6	5.4	586	13	BM191106	da132b10
41	55.6	5.3	569	17	AQ580359	TI36000b
42	50.2	4.8	151	10	BE022923	sm89c11.y
43	47	4.5	711	13	Bj169253	Bj169253
44	46.6	4.4	620	9	AL647838	AL647838
45	45.8	4.4	161	17	B24158	F19B10TR IG

ALIGNMENTS

RESULT 1
AU226180/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AU226180 435 bp mRNA linear EST 23-APR-2002
AU226180 RAF14 Arabidopsis thaliana cDNA clone RAF14-14-A20 3',
mRNA sequence.
AU226180
AU226180.1 GI:19740827
EST.
thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 435)
Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,
Akiyama,K., Enju,A., Oono,Y., Sakurai,T., Carninci,P., Kawai,J.,
Itoh,M., Ishii,Y., Arakawa,T., Shibata,K., Shingawa,A., Muramatsu
M., Hayashizaki,Y. and Shinozaki,K.
Large scale analysis of Arabidopsis full-length cDNA
Unpublished (2002)
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rkc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). cDNA cleaved with BamHI

and XhoI was ligated to modified Lambda FLX-1 vector (Carninci et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified pBluescript vector. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.

FEATURES

source

Location/Qualifiers

1.435
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="RAF14-14-A20"
/clone_1ib="RAF14"
/tissue_type="root"
/lab_host="DH10B"
/note="Site_1: BamHI; Site_2: SalI"

BASE COUNT 117 a 85 c 85 g 148 t
ORIGIN

Query Match 27.5%; Score 288.8; DB 9; Length 435;
Best Local Similarity 90.3%; Pred. No. 2.2e-61;

Matches 327; Conservative 0; Mismatches 17; Indels 18; Gaps 1;

QY 706 GAGATAAAAATGGGCTTGGCTCCTACTACTGAGACAGTGTTCGCAAGAGCTCTGA 765

Db 435 GAGATAAAAATGGGCTTGGCTCCTACTACTGAGACAGTGTTCGCAAGAGCTCTGA 376

QY 766 AGCAAAAGCAATTCAGGCTTGGGAACTGGAGATTTCCTCTGCAAGACAGCACTTAC 825

Db 375 AGCAAAAGCAATTCAGGCTTGGGAACTGGGAGGTTTCATCCTATCAAGACAGCACTTAC 316

QY 826 AATAGCAGCAAGCGATGCTTATGCTTCAATGCGATCTTACAAAGGTTCTTAAGGACCTTC 885

Db 315 AATAGCAGCAAGCGATGCTTATGCTTCAATGCGATCTTACAAAGGTTCTTAAGGACCTTC 256

QY 886 CTGATGCTGTCAGTGCCTATTAACGTAAGGAGGAGGAGGCTTAAGGTTAGCGCTTAACCC 945

Db 255 CTGATGCTGTCAGTGCCTATTAACGTAAGGAGGAGGAGGCTTAAGGTTAGCGCTTAACCC 196

QY 946 AAGAGTTAGCAATCAATGATGATGATACACCTAATCTAGTCAAGTAGATCAATTTCT-- 1002

Db 195 AAGAGTTAGCAATCAATGATGATGATACACCTAATCTAGTCAAGTAGATCAATTTCTTANG 136

QY 1003 -----GTGAATATGTTATCTAGTTCGTCCTTTCACCTTCAGCAAACT 1047

Db 135 GGATTTACAAAGATATTTGAACACAGTATTTAGTTGGGTCATTAATCTTCCAGAACT 76

QY 1048 AG 1049

Db 75 AG 74

RESULT 2
BE659005/c 840 bp mRNA 1linear EST 24-MAY-2001
LOCUS GM700008A10G11 Gm-r1070 Glycine max cDNA clone Gm-r1070-2997 3',
DEFINITION mRNA sequence.

ACCESSION BE659005
VERSION BE659005.1 GI:9984897

KEYWORDS EST.

SOURCE soybean.
ORGANISM Glycine max

REFERENCE 1 (bases 1 to 840)
AUTHORS Vodkin, L., Kelm, P., Shoemaker, R., Retzel, E., Khanna, A., Coryell, V.,

Expanding, J., Rapp, C., Shoop, E., Pardinas, J., Liu, L., and Lewin, H.,

A Functional Genomics Program for Soybean (NSF 9872565)

Unpublished (1999)

Other ESTs: AM460204 corresponding to Gm-cl015-3787 (5')
CONTACT: Vodkin, L.O., PI, A Functional Genomics Program for

Soybean (NSF 9872565)

Lewin, H. A., Director, Keck Center for Comparative and Functional

Genomics
University of Illinois

Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA

Tel: (217) 244-6147

Fax: (217) 333-4582

Email: 1-vodkin@uiuc.edu

This clone is available through: Genome Systems, Inc. 4633 World

Parkway Circle St. Louis, Missouri 63134. For further information

call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)

427-3324 or contact: clones@genomesystems.com or info@genome

systems.com web site: www.genomesystems.com

Seq primer: 5'-TTTCTTTTCTTTTCTTTT(A/C/G)-3'.

Location/Qualifiers

1.840
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="Gm-r1070-2997"
/clone_1ib="Gm-r1070"

/note="The library Gm-r1070 is a sequence-driven, rerecked

set of 9,216 clones selected from cDNA libraries from

various tissues and stages of development of soybean that

represent 2,639 sequences from immature cotyledons, 1,770

from immature seed coats, 3,938 from flowers, and 869

from young pods. The 5' ESTs of the source clones from

the different libraries was used to select singletons, or

a representative of each contig, which were rerecked to

form library Gm-r1070. The cDNA clones of the rerecked

Gm-r1070 library were then sequenced at the 3' end. The

contig analysis to select unique genes was performed by

the Laboratory of Ernest Retzel, Center for Computational

Genomics and Bioinformatics, University of Minnesota,

<http://www.cbc.umn.edu/ResearchProjects/soybean/index.html>

. Rerecking was performed by Genome Systems, St. Louis,

<http://www.genomesystems.com>, and 3' sequencing by the

Keck Center for Comparative and Functional Genomics,

University of Illinois,

<http://www.life.uiuc.edu/diotech/keck.html>. Note: The

corresponding 5' EST from each clone in the Gm-r1070

library is listed in the 'OTHER EST' field. The detailed

information on the source library for each clone can also

be obtained by referring to the Genome Systems clone ID of

the original cDNA library that is also listed under

BASE COUNT 242 a 167 c 158 g 240 t 33 others
ORIGIN

Query Match 26.1%; Score 273.8; DB 10; Length 840;

Best Local Similarity 66.1%; Pred. No. 1.3e-57;

Matches 392; Conservative 0; Mismatches 201; Indels 0; Gaps 0;

QY 300 AGCAATTTCCAGCAATGAGGTTGGGTAGAGTTTGTATGCAAGACTGCTACTGAG 359

Db 727 ATGACATTTGCTGTATATGAGTTTATGTCATTAATTTCTTAATGACGACTTTGATGCT 668

QY 360 GTTGAATAGCGAGCAATGAGCTTAAAGTTCTTGATACCAAGAGAGATGAATGGA 419

Db 667 GTAGAGAAAGCTGCAACAAAGCCTTCAAAATCTCCAGAAAACGACGAGATGATG 608

QY 420 ATAGCTTTTGGCTGGATATTTGATGAGAGCAAGTTTGAAGAAAGGTGTCCTCCG 479

Db 607 CAATCTGCAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 548

QY 480 GGAAGAGTTCGAGCTGTCAGATATGATAGTATGATATGATATGATATGATATGATAT 539

Db 547 GGAAGAGTTCGAGCTGTCAGATATGATAGTATGATATGATATGATATGATATGATAT 488

QY 540 TTTCATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 599

Db 487 ATTCATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 428

QY 600 GTAGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATGAT 659

Db 427 GTTGAAGCTGGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 368

[illegible]

OY	265	CTTATTAACGATTTCCCTCTCCGCCGCGAGCTA-----GGATTTTCCAGCA	315
Dd	144	CTTCTTTCCGCTCTTCCTCCGCTCTTGCTCAATCTAATACAAAGCTTGATATATCCGCCGT	203
OY	316	TGAGCTTTGGTGTAGATTTTGTATAGCAACAGCTCTACTGAGGTGTGATAGCAGCA	375
Dd	204	TGAGATTTTGGAGGTGATATTTTGTATAGCTTTAGCGAAGATGAGGTAGAAAATGGTCGA	263
OY	376	TGCAGCTTATTAAGTCTTGATACCAAGAGAGATGAACTCTGGAATAGCTTTTGTGGCT	435
Dd	264	TGGAGCTTTTAAAAATTTGTGAAATTAAGAAAGAAATGGGTCAAGTTGCTTTAGGGT	323
OY	436	TGGATATTGAGGGAGGACCAAGTTTGAAGAAAGTGTTCTCCGGGGAAAGTTGGAGCTG	495
Dd	324	TTGATATTGAGGGAGGACCCCTCTTCCAAAAAGAAATTTTGCTGGAAAGGCTGGAGTGA	383
OY	496	TCCAGATATGCTAGATAGTATTAATTAATGATGTATCATATTTTTCATCTCGGTATCC	555
Dd	384	TGCAGATATGTTGTGACAGTCAGTATTTGTATGTATGCATATTTTTCATCTGGCAATAC	443
OY	556	CTCAAGTCTCCACACTTCTATTTGAAGATTCACACTTGTAAGGTAGTATTTGAATTG	615
Dd	444	CTCAAGTCTGAGGTTCTCTCTTGTGGGACTCCGAATATATAAAGTTGGAGTTGCAATTG	503
OY	616	ATGGGATCTGTGAAGCTTTCCATGACTATGAGTAGTATTCAAAAGTGTGGAGATC	675
Dd	504	ATGGGATGCTGTCAAGGTCTCAGTCACTATTAAGTATCTTTATAGTTTGGAAAGATC	563
OY	676	TTTTCAGATTTAGCCCAACCAAAAAATTTGGTGAGAT---AAAAAATGCGGCTTGCCTCAC	732
Dd	564	TTTTCAGATCTAGCGAATCAAAAATTTTGGTAGTATTTGCCGCGACGTGAGATCTTGGCTGCTC	623
OY	733	TAACTGAGACACTTGTTTTGCAAGAGCTCTGAGAACCAACGAATCAGGCTTGGGAATC	792
Dd	624	TGACGAGAGAGATTTTATTGGCAAAAGAGCTCCCAAGCCCAAGAAAATTAAGCTGGGAAGCT	683
OY	793	GGGAGTTTATCCCTGTGCAA	813
Dd	684	GGGAGTTATATCTTTATCCA	704

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
RESULT 4 AM201789	552 bp mRNA linear EST 02-DEC-2001	AM201789	AM201789								
1	552 bp mRNA linear EST 02-DEC-2001	AM201789	AM201789								
2	552 bp mRNA linear EST 02-DEC-2001	AM201789	AM201789								
3	552 bp mRNA linear EST 02-DEC-2001	AM201789	AM201789								
4	552 bp mRNA linear EST 02-DEC-2001	AM201789	AM201789								
5	552 bp mRNA linear EST 02-DEC-2001	AM201789	AM201789								
6	552 bp mRNA linear EST 02-DEC-2001	AM201789	AM201789								
7	552 bp mRNA linear EST 02-DEC-2001	AM201789	AM201789								
8	552 bp mRNA linear EST 02-DEC-2001	AM201789	AM201789								
9	552 bp mRNA linear EST 02-DEC-2001	AM201789	AM201789								
10	552 bp mRNA linear EST 02-DEC-2001	AM201789	AM201789								
11	552 bp mRNA linear EST 02-DEC-2001	AM201789	AM201789								
12	552 bp mRNA linear EST 02-DEC-2001	AM201789	AM201789								
13	552 bp mRNA linear EST 02-DEC-2001	AM201789	AM201789								
14	552 bp mRNA linear EST 02-DEC-2001	AM201789	AM201789								
15	552 bp mRNA linear EST 02-DEC-2001	AM201789	AM201789								
16	552 bp mRNA linear EST 02-DEC-2001	AM201789	AM201789								
17	552 bp mRNA linear EST 02-DEC-2001	AM201789	AM201789								
18	552 bp mRNA linear EST 02-DEC-2001	AM201789	AM201789								
19	552 bp mRNA linear EST 02-DEC-2001	AM201789	AM201789								
20	552 bp mRNA linear EST 02-DEC-2001	AM201789	AM201789								
21	552 bp mRNA linear EST 02-DEC-2001	AM201789	AM201789								
22	552 bp mRNA linear EST 02-DEC-2001	AM201789	AM201789								
23	552 bp mRNA linear EST 02-DEC-2001	AM201789	AM201789								
24	552 bp mRNA linear EST 02-DEC-2001	AM201789	AM201789								
25	552 bp mRNA linear EST 02-DEC-2001	AM201789	AM201789								
26	552 bp mRNA linear EST 02-DEC-2001	AM201789	AM201789								
27	552 bp mRNA linear EST 02-DEC-2001	AM201789	AM201789								
28	552 bp mRNA linear EST 02-DEC-2001	AM201789	AM201789								
29	552 bp mRNA linear EST 02-DEC-2001	AM201789	AM201789								
30	552 bp mRNA linear EST 02-DEC-2001	AM201789	AM201789								
31	552 bp mRNA linear EST 02-DEC-2001	AM201789	AM201789								
32	552 bp mRNA linear EST 02-DEC-2001	AM201789	AM201789								
33	552 bp mRNA linear EST 02-DEC-2001	AM201789	AM201789								
34	552 bp mRNA linear EST 02-DEC-2001	AM201789	AM201789								
35	552 bp mRNA linear EST 02-DEC-2001	AM201789	AM201789								
36	552 bp mRNA linear EST 02-DEC-2001	AM201789	AM201789								
37	552 bp mRNA linear EST 02-DEC-2001	AM201789	AM201789								
38	552 bp mRNA linear EST 02-DEC-2										

Db 515 TTTCAGATTAGCCACCAAAAATGTTGGAGATAAAAATGGGCGCTTCCTACTAA 574
Oy 736 CTGAGACCTGTTTGGAAAGAGCTCTGAGCCAAACAGATCAGCTTGGGAACGGG 795
Db 575 CTGAGACCTGTTTGGAAAGAGCTCTGAGCCAAACAGATCAGCTTGGGAACGGG 634
Oy 796 AGTTTATCCTCTGCAAGCAGCAGTTACATACGC 832
Db 635 AGGTGATNCTCTATCAAGCAGCAGTTACATACGC 671

RESULT 6
AM832139
LOCUS
DEFINITION
AM832139 467 bp mRNA linear EST 03-DEC-2001
sm20b11.y1 Gm-c1027 glycine max cdna clone GENOME SYSTEMS CLONE ID:
Gm-c1027-9430 5' similar to SW:WRN_HUMAN Q14191 WERNER SYNDROME
HEITCASE. [1] ; mRNA sequence.
AM832139
AM832139.1 GI:7926113
EST.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
glycine max
soybean.
Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 467)
Shoemaker, R., Keim, P., Vodkin, L., Erpeiding, J., Coryell, V., Khanna
, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
Wyle, T., Underwood, K., Stepien, M., Theissen, B., Allen, M., Bowers
, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harey, N., Schurk
, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann
, R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Resgen, Invitrogen Corp. 2130
South Memorial Parkway Huntville, AL 35801 For further information
call: (800)-533-4363 or contact via email: cou@resgen.com
Insert Length: 734 Std Error: 0.00
High quality sequence stop: 417.
Location/Qualifiers
1. 467
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1027-9430"
/clone_id="Gm-c1027"
/tissue_type="cotyledons of 3- and 7-day-old Williams
seedlings"
/lab_host="DH10B"
/note="Vector: pBluescript II SK+, Site_1: EcoRI; Site_2:
XhoI; This cDNA library was constructed from mRNA isolated
from cotyledons of 3- and 7-day-old Williams seedlings
which were propagated on paper towels with distilled
water. The cotyledons were flash-frozen in liquid
nitrogen, then lyophilized for 72 hours. Unequal amounts
of mRNA was used for cDNA synthesis. Stratagene's cDNA
Synthesis Kit (catalog number 200401) was used to
synthesize the cDNA. First-strand synthesis was
performed with 5-methyl dCTP, hence the ligated cDNA was
hemimethylated. A modification of Stratagene's
first-strand synthesis primer was used. An anchor
nucleotide (V-A, C, or G) was added to the 3' end of the
primer. [GAGAGAGAGAGAGAGAGAGAGAGCTGAGT]18 to anchor
the primer at the 5' end of the poly(A) tract. After
second-strand synthesis, the cDNA ends were filled in
with cloned Pfu DNA, ligated to EcoRI adapters and

subsequently phosphorylated. The XhoI site within the
first-strand synthesis primer was then restricted by
digestion with XhoI; all XhoI sites in the cDNA would be
protected by their hemimethylated status. The cDNA
constructs were size-fractionated with a 500 bp cutoff,
using GIBCOBRL Life Technologies' cDNA Size Fractionation
column. The column eluent was then ligated into
Stratagene's pBluescript(tm) II XR predigested vector
(pBluescript II SK(+)) that has been digested with EcoRI
and XhoI, and phosphorylated by Stratagene). 97% of the
white and blue colonies appear to contain recombinant
plasmids with cDNA inserts, based on size (n=30). This
library was constructed by Dr. Paul Keim and Dr. Virginia
Coryell."

BASE COUNT 133 a 84 c 109 g 140 t 1 others
ORIGIN

Query Match 23.0%; Score 241.2; DB 10; Length 467;
Best Local Similarity 70.0%; Pred. No. 1.6e-49;
Matches 324; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

Oy 430 TTGGCTTGATTTGAGTGGAGACCAAGTTTACAAAAGTGTCTCCCGGGAAGCTTG 489
Db 4 TTGGATTTGACATTTGAGTGAAGAACCCACCTTCAGAAAAGTGTCTCCCGGGAAGCTTG 63
Oy 490 CGAGCTGCAGATATGTGTAGATAGTATTTATTTGAGATGATAGATTTTTCATCTCG 549
Db 64 CAGTGATGCAATATATGTGTAGTACACTACATTTGATTTGATTTTATTTATTTCTG 123
Oy 550 GTATCCCTCAAGTGTCCCAACATCTTATTTGAAGATTTCAACTGTGAAGTAGTATTTG 609
Db 124 GAATCCCTCAAAATTTACAGCTTTTGTCTGAAGATCCACAGCTCTGAAGTTGAGCTG 183
Oy 610 GAATGATGTTGACTGTGTGAAGCTTTTTCATGACTATGATGATGATGATGATGATG 669
Db 184 GGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 243
Oy 670 AGGATCTTTCAGATTTTACCAACCAAAATTTGTGTGAAGATGATGATGATGATGATG 729
Db 244 CGGATCTTCTTTCATCTATATCAAAAGCTGTGTGTGAAGATGATGATGATGATGATG 303
Oy 730 CACTACACTGACACTGTTTTCGCAAAAGAGCTCTGCAAGCCAAACAGATTCAGGCTTG 789
Db 304 CTTTGACCTGAAAACCTTATTCAAAACAGCTTAAACCAACAAATTAAGTGGGAA 363
Oy 790 ACTGGAGTTTATTCCTGTGTCAAGCAGCTTACATTAAGCAGCAAGAGATGCTATG 849
Db 364 ATGGGAGGCTCTGTTTGTCAAAAGAGCAACTGATGATGATGATGATGATGATGATG 423
Oy 850 CTTGATGCACTTTTACAGGTTCTTAGAGACCTTCGATGAGC 892
Db 424 CTTTGTGTGTTTATTCAGGCGATTAAGATTCCTCCGAGGC 466

RESULT 7
BO799068
LOCUS
DEFINITION
BO799068 661 bp mRNA linear EST 30-JUL-2002
EST 1237 Green Grape berries lambda Zap II library Vitis vinifera
cDNA clone GI172E01 3', mRNA sequence.
BO799068
BO799068.1 GI:22014034
EST.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Vitis vinifera.
Vitis vinifera
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Vitaceae; Vitis.
1 (bases 1 to 661)
Abdal, P., Agasse, A., Ageorges, A., Anassova, R., Barrieu, F.,
Couture, C., Dedaldecamp, F., Delrot, S., Glissant, D., Grimpel, J.,
Hamdi, S., Romieu, C. and Terrier, N.
Generation of Expressed Sequence Tag from Grape Berry (skin, pulp
or seeds) at Various Developmental Stages

[illegible]

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
723	CTTGGCCTACACAACTGAGACCTGTTTGGTCCAAAGAGCTCCGAAAGCCAAACAGTAATAGG	782		
723	CTTGGCCTACACAACTGAGACCTGTTTGGTCCAAAGAGCTCCGAAAGCCAAACAGTAATAGG	782		
362	CTTACATCATTTAACTGAGACCTTCTTCCCAAGCAGCTCCCAAGAGCAAGTAATATACGG	421		
783	CTTGGGAACAGCTGGAGTTTATCTCTGTCCAAGCAGCAGTTACATATACGACCAACGGAT	842		
422	TTGGGGGAATTTGGGAGGCTAATGTGTATATCTAGGACCAACATATATATGCTATACAGAT	481		
843	GCTTATATGCTTCATGGCATCTTTTA	865		
482	GCCTTTGTTCCTGCTGTAATATA	504		
RESULT 10				
BM094501				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
COMMENT				
FEATURES				
SOURCE				

ORIGIN	Query Match	20.4%	Score 213.6	DB 13	Length 531
	Best Local Similarity	64.6%	0	Pred. No. 1.2e-42	
	Matches 340	Conservative	0	Mismatches 174	Indels 12; Gaps 1
QY	299 TAGGAATTTTCACCAATGAGTTTGGCGTAGAGATTTTGTATGACAGACTGCTACGA	358			
DB	18 TATGAGATGGCTCTAATGAGTTTGTAGTGGTCAATTTTATATACAGACTTTTGTATGC	77			
QY	359 GGTGATATAGCAGACATGACGCTTATTTAAAGTTCTTGATATCCAGACAGATGAATCTGG	418			
DB	78 TGTAGAGAAACCTGCACAAAGGCTCTTACAAATTCCTCCAGAGATGATGACA-----	128			
QY	419 AATAGCTTTTGTGGCTGTGATTTATGATGAGAGACCAAGTTTGTGAAAAGGTTCTGCC	478			
DB	129 ---AATTCATTTGTGATTTGTGACATTTGATGATGGAACCCACTTCAGAAAAGTGTCCACC	185			
QY	479 GGGGAAGGTTGCGACTGCTCCAGATATGTGTAGATAGTAATTAATTTGATGTTATGATAT	538			
DB	186 CGGAAGAGTACAGATGATGACATATATAGTGACACTGATGATCTGATCTTCTTACATCT	245			
QY	539 TTTTCATCTGCTGATTCCTCCAAAGTCCCAACATCTTTATGAAGTTCCACATCTGTAA	598			
DB	246 AATTCATCTGGAATCCCTGGAATTTACACTTTTGCTTGAAGATCCACAGCTTTGAA	305			
QY	599 GGTAGATATTTGGAATTTGATGCTGACTGCTGAAGCTTTTCCATGACTATGAGTTAGTAT	658			
DB	306 GGTGGAGCTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	365			
QY	659 CAAAGATTTAGCATCTTTAGATTTTACCAACCAAAAATTTGGTGAGATTAATAATG	718			
DB	366 TAAAGTGTGAGGATCTTTCTTTTCATGATGATGATGATGATGATGATGATGATGATGAT	425			
QY	719 GGGCTTCCCTACACTAACTGAGACACTGTTGTCGAAAGAGTCCGGAAGCCAAACAGAT	778			
DB	426 GGGCTTTCATCTTTGACTGAAAACCTTTTATCAAAACAGCTTAAAGAGCTTAACAAAT	485			
QY	779 CAGGCTTGGAGACTGGAGTTTATCTCTGTCAAAAGCAGAGTTA	824			
DB	486 AAGACTGGGAATTTGGGAGACTCCTGTTTGTCAAAAGAGCACTA	531			
RESULT 11					
BM358892		515 bp	mRNA	linear	EST 09-JAN-2002
LOCUS	GA_Ea0013P02r	Gossypium arboreum	7-10 dpa	fiber	library Gossypium
DEFINITION	arboresum cDNA clone GA_Ea0013P02r,				mRNA sequence.
ACCESSION	BM358892				
VERSION	BM358892.1	GI:18099638			
KEYWORDS	EST.				
SOURCE	Gossypium arboreum.				
ORGANISM	Gossypium arboreum				
REFERENCE	Eukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
AUTHORS	Rosidae; eudicots II; Malvales; Malvaceae; Gossypium.				
TITLE	1 (bases 1 to 515)				
JOURNAL	Wing,R.A., Frisch,D., Yu,Y., Mait,D., Rambo,T., Simmons,J., Henry				
COMMENT	D., Wood,T.C., Leslie,A. and Wilkins,T.A.				
	An integrated analysis of the genetics, development, and evolution				
	of the cotton fiber				
	Unpublished (2000)				
	Contact: Wing RA				
	Clemson University Genomics Institute				
	Clemson University				
	100 Jordan Hall, Clemson, SC 29634, USA				
	Tel: 864 656 7288				
	Fax: 864 656 4293				
	Email: rwing@clemson.edu				
	Total High Quality bases = 415				
	Seq primer: TAAATGACATCACTATATGGG				
	High quality sequence stop: 514.				
	Location/Qualifiers				

	source	1..515 /organism="Gossypium arboreum" /strain="AKA" /cultivar="8400" /db_xref="taxon:29729" /clone="GA_Ea0013P02r" /clone_id="Gossypium arboreum 7-10 dpa fiber library" /tissue_type="Fibers isolated from bolls harvested 7-10 dpa" /lab_host="E. coli" /note="vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT	134 a 112 c 109 g 160 t	
ORIGIN		
Query Match	16.4% Score 172.2; DB 13; Length 515;	
Best Local Similarity	64.6%; Pred. No. 2.6e-32;	
Matches 277; Conservative	0; Mismatches 145; Indels 9; Gaps 1;	
OY	208 ATCAATCCCATTAATAATATCCTCGGCCAAATTGCCTGTTCATCATCTTTTACATCTT	267
Dd	 ATCAAGATATACCCTCAAACAACGCGCCAATTGGCCCCATTTCTATCTCACCCCCTT	146
OY	268 APTAACGATTTCTCTCTCCCGCTTGCGAGCTA-----GGAAATTTCCAGCAATGA	318
Dd	 CTTTTCGCTCTCTCGCTCGCTGTGTCAATCTAATACAGAGCTTAGATATCCGCCCTTGA	206
OY	319 GGTTTTGGTGTAGATTTTGTATATGACCAAGACTGCTACTGAGTGTATTAAGCAGAATGC	378
Dd	 GATTTGGAGGTGCATATTTTGTATATGACTTACGGAAGATGAGAGTGAAGAAAATGCTGCAATGG	266
OY	379 AGCTTATTAAGTTCTTGATATACCAAGACATGAATCTGGAATAGCTTTTGTGGCTTGG	438
Dd	 AGCTATTAAAAATTTGTGTAAATTTAAAGAAAGAAATGGGTCAAGTTCTTTAGGGTTTG	326
OY	439 ATATTGAGTGGAGACCAAGTTTATGAAAGAGTGTTCTCCCGGGAGAGTTGCGACTGCC	498
Dd	 ATATTGAGTGGAGAGCCCTCTTCCAAAAGGAATTTTCCCTGGAGAGCTCGGTGATGC	386
OY	499 AGATATGTGTAGATATTAATTAATGTCGATGTTATGATATTTTTCATTCGTGATCCCTC	558
Dd	 AGATATGTGTGACAGCTACGATATGTTATGATGATGATATTTTCATCTCGCATATACCTC	446
OY	559 AAAGCTCCCAACATCTTTATGAAGTTCAACACTGTGAAGGTATTTGGAATGTGATG	618
Dd	 AAAGCTCCACAGGTTCTTTGAGGACCTCGAAATTAATAAAGTTGGAAGTTCAATGTATG	506
OY	619 GTGACTCTG 627	
Dd	 GGGATGCTG 515	
RESULT 12		
LOCUS	BM358891	
DEFINITION	517 bp mRNA linear EST 09-JAN-2002	
ACCESSION	GA_Ea0013O22r Gossypium arboreum 7-10 dpa fiber library	
VERSION	arboeum cdna clone GA_Ea0013O22r, mRNA sequence.	
KEYWORDS	BM358881 BM358891 GI:18099637	
SOURCE	EST.	
ORGANISM	Gossypium arboreum. Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.	
REFERENCE	Wing,R.A., Fritsch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry	
AUTHORS	,D., Wood,T.C., Leslie,A. and Wilkins,T.A.	
TITLE	An integrated analysis of the genetics, development, and evolution	
JOURNAL	of the cotton fiber unpublished (2000)	
COMMENT	Contact: Wing RA Clemson University Genomics Institute Clemson University	

Oy	531	ATGCAATATTTTCATTCGCGATGATCCCTCAAGTCCCAAGTCTTAATGAAGATTCAACA	590
Dd	406	CTACATCTAATTCATTCGCGATGATCCCTCAAAATTACAGCTTTGCTTGAAGATCCCA	465
Oy	591	CTTGTAAGTAGTAGTATTGTAAGTATGATGCTGACTCTGTGAAGCTTTTCATGACTATGGA	650
Dd	466	GCTTTGAAGAGTTGGAGCTGGAGGTGATGATGCTGTGAAGGTTTATAGAGATTATTAAC	525
Oy	651	GTTAGTATCAACAAGATGTTGAGATCTTTC	679
Dd	526	ATATCTGTTTAAAGGTGATGACGATCTTTC	554
RESULT 14			
LOCUS	BJ213300	595 bp	mRNA
DEFINITION	BJ213300 Y. Ogihara unpublished cDNA library, wh Triticum aestivum		
ACCESSION	CDNA clone wh21m08 5', mRNA sequence.		
VERSION	BJ213300		
KEYWORDS	BJ213300.1	GI:19952807	
SOURCE	EST.		
ORGANISM	bread wheat.		
REFERENCE	Triticum aestivum		
AUTHORS	Eukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
TITLE	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidaeae		
JOURNAL	Triticaceae; Triticum.		
COMMENT	1 (bases 1 to 595).		
	Ogihara,Y and Murai,K.		
	Expressed genes in Triticum aestivum		
	Unpublished (2002)		
	Contact: Tadasu Shin-1		
	Center For Genetic Resource Information		
	National Institute of Genetics		
	1111 Yata, Mishima, Shizuoka 411-8540, Japan		
	Tel: 81-559-81-6856		
	Fax: 81-559-81-6855		
	Email: tshin1@genes.nig.ac.jp.		
FEATURES	Location/Qualifiers		
source	1..595		
	/organism="Triticum aestivum"		
	/cultivar="Chinese Spring"		
	/db_xref="taxon:4565"		
	/clone="wh21m08"		
	/clone_1ib="Y. Ogihara unpublished cDNA library, wh"		
	/tissue.type="spike at meiosis"		
	/dev_stage="Peekes' scale 9"		
	/note="Vector: lambda uni-ZAP XR, excised phagemid;		
	Site_1: EcoRI; Site_2: XhoI; Plants were grown under		
	hydroponic conditions at UC Davis, salt stressed for 12		
	hours, and for 7 days, then dissected and frozen (Akhunov		
	in J Dvorak lab). Total RNA was prepared from sheath		
	tissue, equal quantities of RNA were pooled from the two		
	samples, polyA was purified from the pooled RNA, a cDNA		
	library was made, and the cDNA clones were in vivo		
	excised to give pluescript phagemids in the TV Close lab		
	at the University of California, Riverside (Akhunov, Chinn,		
	Choi, Close, Fenton, Kianian, Otto, Simons, zhang).		
	Plasmid DNA preparations and DNA sequencing were		
	performed in the OD Anderson lab (all other authors)."		
BASE COUNT	170 a 119 c 148 g 157 t		
ORIGIN	1 others		
Query Match	15.4%, Score 162; DB 13; Length 595;		
Best Local Similarity	58.3%, Pred. No. 9,le=30;		
Matches 323; Conservative	0; Mismatches 225; Indels 6; Gaps 2;		
Oy	321	TTTGGGTAGGATTTTGTATATAGCAAGACTCTACTGAGTTGATTAACGAGCAATGCAG	380
Dd	45	TTCAGTGGCAAGATAGTACTACGCGGACGCAATTGAATGGAGAAAGTACAGGGGAG	104
Oy	381	CTTATTAAGTCTTGATACCAAGAGATTAATCTGGAATAGCTTTTGTGGCTTGAT	440

Db		105	ATTGTAGCGAAAATTCGAAAGCATGA---AGGCCCTCTGGCCCGGCTCCTTGGTTGCAT	161			
Oy		441	ATTAGTGAGACCAACTTTTAGAAAAGGTGTTCCTCCGGGAAAGGTGGACGTGTCCAG	500			
Db		162	CTCAGAGTGAACCCCTTCCAGNAGAGGAAACCACATGTAAAGTCGGGTGATCCAA	221			
Oy		501	ATAATGTAGATAGTAATTAATTTGATGATTAATGATATTTTCAATCTGTATCCCTCAA	560			
Db		222	TTATGCATGAGCAAACTCATTTGTATCTCATGATATTCATCTGTGGGTGCTCCC	281			
Oy		561	AGTCTCAACACTTTAATGAAGATTCAACACTGTAAAGTAGTATTTGCAATTGATGT	620			
Db		282	ATCTTGAANAATCTCTTTTGGAGGACAGTTCATCCCTTAAAGTTGGAGTATATGACAAAT	341			
Oy		621	GACTCTGTGAAGCTTTCCATGACTATGATGATTTGATATCAAGAATGTTGGAGACTTCCA	680			
Db		342	GATCAAGAAGAAATGTTCAATGATTAATGATGTCCGTGTACAAACATTTGATGATTTTCA	401			
Oy		681	GATTTAGCCAACCAAAAAATTTGTGTG--AGATAAAAATGGGGCTTGGCTCACTAACT	737			
Db		402	ACTGTGCAAAACGTCAAGTTAGTGTGGCCCTATPAAAGATGAGATCTTGCTGCTTAAT	461			
Oy		738	GAGCACTTGTTCGCAAGAGCTCTCTGAAGCCAAACAGAAATACAGCTTGGAACTGGGAG	797			
Db		462	GAAATGGTTACATGTAAAGAGTGTCCAAAGCTGGCAACATTAACAATGGGAAATGGGAG	521			
Oy		798	TTTTATTCCTGTGTAAGACAGCACTTACAAATACGACCAACGATGCTTTAGCTTTCATG	857			
Db		522	TCTTTGTGTCTCTCAAAAAAGCACTTGATGATGTCTGTACCGATGCTCATATCTGTGG	581			
Oy		858	CATCTTTTACAAGT 871				
Db		582	TACTTGTATGAGT 595				
<hr/>							
RESULT_15							
B1433038							
LOCUS	B1433038	499 bp	mRNA	linear EST 21-AUG-2001			
DEFINITION	EST535799 P. infestans-challenged leaf Solanum tuberosum cDNA clone						
ACCESSION	PC04286.5	sequence, mRNA sequence.					
VERSION	B1433038						
KEYWORDS	B1433038.1 GI:15257728						
SOURCE	EST.						
ORGANISM	potato.						
	Solanum tuberosum						
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;						
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;						
	Asteridae; euasterids I; Solanales; Solanaceae; Solanum.						
REFERENCE	1 (bases 1 to 499)						
AUTHORS	Restrepo-O., Griffiths,H.M., Smart,C.D., Cho,J., Chelmingo,A.,						
	Bougrit,O., Buell,C.R., Rohnig,C.M., Fry,M.E. and Baker,B.						
	Generation of ESTs from potato leaves Challenged with Phytophthora						
	infestans, Compatible Interaction						
TITLE	Unpublished (2000)						
JOURNAL	Contact: Cathy Rohnig						
COMMENT	The Institute for Genomic Research						
	For clone info: please contact Research Genetics, Libraries						
	Division tel 1-800-711-6195, email cdnare@sgen.com						
	Seq primer: M13P-R.						
FEATURES	Location/Qualifiers						
source	I..499						
	/organism="Solanum tuberosum"						
	/cultivar="Kennebec"						
	/db_xref="taxon:4113"						
	/clone="PPC4286"						
	/clone_id="P. infestans-challenged leaf"						
	/tissue_type="leaf"						
	/dev_stage="6 week old"						
	/lab_host="SOLR"						
	/note="Vector: pliuscript SK(-); Site 1: EcorI; Site_2:						
	XhoI; Whole plants were challenged with 450,000						
	sporangia/ml P. infestans US-1(US 940501) in Biotron						
	(Madison, Wisconsin). Leaf tissue was collected at 1, 2,						

5, 12, and 24 hours post-challenge and frozen in liquid nitrogen immediately upon removal. Kennebec plants showed no signs of HR. Katahdin plants (susceptible to P. infestans US-1) were used as controls and showed infection. NOTE: We cannot exclude the possibility that this sequence is actually derived from Phytophthora rather than potato."

BASE COUNT 141 a 86 c 118 g 154 t
ORIGIN

Query Match 14.9%; Score 156.8; DB 13; Length 499;
Best Local Similarity 62.5%; Pred. No. 1.8e-28;
Matches 243; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

```
OY 296 AGCTAGCAATTTCCAGCAATGAGTGTGGTGTAGATTTGTATAGCAGACTGCTAC 355
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 108 AGTATGCAATACCCGGAATCTCTTTAAAGACATGTATATACAGCACTATTAA 167
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 356 TGAGTTGATTAAGCAGCAATGCACTTATTAAAGTCTTGATACCAAGAGATGAATC 415
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 168 AGAGGTGGAATACTGCAGTAGAGCTGTAAATTTTGTGAGAAAAGAGAGAAAGA 227
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 416 TGGATAGCTTTGTGGCTTGATATTGAGTGAACCACTTTAGAAAAGTGTCT 475
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 228 AGGGAATGTGCTTTGATTTGACATTGAATGAAGCCCACTTTAGAAAGAGTGTGCC 287
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 476 CCCGGGAAGGTGGGACGTGCCAGATATGATAGTATATTTGTGATGTATGCA 535
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 288 ACCTGGGAAGCCCGTGTATAGCAGATATGTGTACAAAGGTAATGTATATGTCTGCA 347
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 536 TATTTTCATTCCTGTATCCCTCAAGTCTCCACATCTTATTGAAGATTCACACTGT 595
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 348 TATCATCCACTGTGAATCCCTCAAACTGTGAATCTCTTGTGAGATCCAACTGTGT 407
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 596 AAAGTACGTATGATGATGAGTGAAGTCTGTGAAGCTTTCCATGACTATGAGTTAG 655
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 408 GAAGGTGGGTGTTCATTCGCAAAAGATGCTTACAAAGTTCGCAAGATCCACATGTATC 467
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 656 TATCAAGATGTGAGATCTTCAGATTTAG 687
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 468 TGTGAAGGCTTTGGAAGATCTTCTGAACCTTG 499
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: June 14, 2003, 08:23:16
Job time : 1830 secs

396 GATACCAAGAGAGATGAATCTGGAATAGCTTTTGTGTTGGCTTGATATTGAGTGAGAGACCA 455

```
Db 433 GATATTGACATGACTCTATCAGATGGGAGTGTGGGATTTGACATGAGTCCACCA 492
Qy 456 AGTTTGAAGAAAGTCTTCCTCCCGGAGAGTTGCGACTGTCAGATATGTGATAGT 515
Db 493 TTATACAAATAGAGGGAACCT---TGCAAAAGTTGCACTAATTCAGTTGTGTGTTCCAG 549
Qy 516 AATTATTGTGATGTTATGCAATATTTT---CATTCGTGATCCCTCAAAAGTCTCCAAAT 572
Db 550 AGCAAAATGTTACTGTTCCACAGCTTTCCTCCATGTCAGTATTTTCCCGAGGATTTAAATG 609
Qy 573 CTTATTGAGATCAACCTGTGTAAAGTAGTATGATGATTTGAAATTTGATGACTGCTGAG 632
Db 610 TTGCTTAAATTAAGACAGTTTAAAGGCGAGCTGTAGAAATTTGAGAGATGACTGGAAA 669
Qy 633 CTTTTCATGACATGAGTGTAGTATCAAGATGTTGAGATCTTTCAGATTTTACCCAAC 692
Db 670 CTTCTACGTGACTTGTATCAATTAAGAAATTTTGGAGTTGACAGATGTTGCCAAT 729
Qy 693 CAAAAAATTGGTGGAGTAAAAAATGGGCGCTTCCCTCCTACTGACAGACATTTGTTGC 752
Db 730 AAAAGCTGAATGTACAGAGACCTGAGCCTTAACTGCTGTTAAACACCTCTTAGT 789
Qy 753 AAAGAGCTCCTGAAGCCAAAGAAATGAGCTTGGAACTGGAGTTTATCCTCTGTCA 812
Db 790 AAACAGCTCCTGAAGACAAAGTCTATCCGCTGTAGCAATGGAGTAAATTTCTCTCACT 849
Qy 813 AAGCAGAGTTTACATACGAGCAACGAGATGCTTATGCTTACGATGCACTTTACAGGTT 872
Db 850 GAGACCGAAGAACTGTATGACAGCCACATGATGCTTATGCTGTTTATTTATTTACGAAAT 909
Qy 873 CTTAAGGACCTTCTGATGCTGT 895
Db 910 TTAGAGATTTTGATGATGACTGT 932

RESULT 2
US-09-791-211-11
; Sequence 11, Application US/09791211
; Patent No. 6448080
; GENERAL INFORMATION:
; APPLICANT: Donna T. Ward
; APPLICANT: Andrew T. Watt
; TITLE OF INVENTION: ANTISENSE MODULATION OF WRN EXPRESSION
; FILE REFERENCE: RTS-0205
; CURRENT APPLICATION NUMBER: US/09/791, 211
; CURRENT FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 90
; SEQ ID NO 11
; LENGTH: 5208
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (232)...(4530)
US-09-791-211-11
```

Query Match 10.6%; Score 111; DB 4; Length 5208;
Best Local Similarity 54.1%; Pred. No. 2.5e-25;
Matches 272; Conservative 0; Mismatches 225; Indels 6; Gaps 2;

```
Qy 396 GATACCAAGAGATGATGATGCAATAGCTTTTGTGGCTTGGATATGAGTGGAGACCA 455
Db 433 GATATTGACATGACTCTATCAGATGGGAGTGTGGGATTTGACATGAGTGGCCACCA 492
Qy 456 AGTTTGAAGAAAGTCTTCCTCCCGGAGAGTTGCGACTGTCAGATATGTGATAGT 515
Db 493 TTATACAAATAGAGGGAACCT---TGCAAAAGTTGCACTAATTCAGTTGTGTGTTCCAG 549
Qy 516 AATTATTGTGATGTTATGCAATATTTT---CATTCGTGATCCCTCAAAAGTCTCCAAAT 572
Db 550 AGCAAAATGTTACTGTTCCACAGCTTTCCTCCATGTCAGTATTTTCCCGAGGATTTAAATG 609
Qy 573 CTTATTGAGATCAACCTGTGTAAAGTAGTATGATGATTTGAAATTTGATGACTGCTGAG 632
```

```
Db 610 TTGCTTAAATTAAGACAGTTTAAAGGCGAGTGTAGAAATTTGAGAGATCACTGGAAA 669
Qy 633 CTTTTCATGACATGAGTGTAGTATCAAGATGTTGAGATCTTTGACATTTAGCCAAC 692
Db 670 CTTCTACGTGACTTGTATCAATTAAGAAATTTTGGAGTTGACAGATGTTGCCAAT 729
Qy 693 CAAAAAATTGGTGGAGTAAAAAATGGGCGCTTCCCTCCTACTGACAGACATTTGTTGC 752
Db 730 AAAAGCTGAATGTACAGAGACCTGAGCCTTAACTGCTGTTAAACACCTCTTAGT 789
Qy 753 AAAGAGCTCCTGAAGCCAAAGAAATGAGCTTGGAACTGGAGTTTATCCTCTGTCA 812
Db 790 AAACAGCTCCTGAAGACAAAGTCTATCCGCTGTAGCAATGGAGTAAATTTCTCTCACT 849
Qy 813 AAGCAGAGTTTACATACGAGCAACGAGATGCTTATGCTTACGATGCACTTTACAGGTT 872
Db 850 GAGACCGAAGAACTGTATGACAGCCACATGATGCTTATGCTGTTTATTTATTTACGAAAT 909
Qy 873 CTTAAGGACCTTCTGATGCTGT 895
Db 910 TTAGAGATTTTGATGATGACTGT 932

RESULT 3
US-09-127-670-5
; Sequence 5, Application US/09127670
; Patent No. 6228583
; GENERAL INFORMATION:
; APPLICANT: Massachusetts Institute for Technology
; APPLICANT: Leonard P. Guarente
; APPLICANT: David A. Sinclair
; APPLICANT: David B. Lombard
; TITLE OF INVENTION: ASSAYS FOR COMPOUNDS WHICH EXTEND LIFE
; FILE REFERENCE: MIT-7720PA
; CURRENT APPLICATION NUMBER: US/09/127, 670
; CURRENT FILING DATE: 1998-07-31
; EARLIER APPLICATION NUMBER: US 60/054, 629
; EARLIER FILING DATE: 1997-08-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PastedSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 6476
; TYPE: DNA
; ORGANISM: Murine
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (229)...(4432)
US-09-127-670-5
```

Query Match 7.9%; Score 82.8; DB 4; Length 6476;
Best Local Similarity 53.2%; Pred. No. 3.7e-16;
Matches 199; Conservative 0; Mismatches 172; Indels 3; Gaps 1;

```
Qy 481 GGAAGTTGCGACTGTCCAGATATGTGTAGATAGTAAATTTGATGTTATGCAATTTT 540
Db 494 GCAAGTGCAGATGATGATGATGCTGTGCTGTGAGCAAAATTTACTTTCACATTTT 553
Qy 541 TT---CATTCGTGATCCCTCAAAAGTCTCCAACTCTTATTTGAAGATTCAACTGTAA 597
Db 554 CTTCCATGATGATGTTTCCCGAGGATTTAAATGTTATGAGAAACAAATTCATTTAGA 613
Qy 598 AGTTAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 657
Db 614 AGGAGAGGAGTTGGATTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 673
Qy 658 TCAAGATGTTGAGATCTTTCAGATTTTACCAACCAAAATTTGTTGAGATTTAAAT 717
Db 674 TGAGAGATTTTGTGAGACTGACGAGATGTTGCAATTTGAAAGTTGAGAGTGGAGACT 733
Qy 718 GGGGCTTGGCTCTACTGACAGACATTTGTTGCAAAAGAGCTTCTTAACCAACGAA 777
```



```
RESULT 6
US-08-781-891-207
: Sequence 207, Application US/08781891
: Patent No. 6090520
: GENERAL INFORMATION:
:   APPLICANT: Fu, Ying-Hui
:   APPLICANT: Yu, Chang-En
:   APPLICANT: Oshima, Junko
:   APPLICANT: Mulligan, John T.
:   TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
:   NUMBER OF SEQUENCES: 209
:   CORRESPONDENCE ADDRESS:
:     ADDRESSEE: SEED and BERRY LLP
:     STREET: 6300 Columbia Center, 701 Fifth Avenue
:     CITY: Seattle
:     STATE: Washington
:     COUNTRY: USA
:     ZIP: 98104-7092
:   COMPUTER READABLE FORM:
:     MEDIUM TYPE: Floppy disk
:     COMPUTER: IBM PC compatible
:     OPERATING SYSTEM: PC-DOS/MS-DOS
:     SOFTWARE: Patentin Release #1.0, Version #1.30
:     CURRENT APPLICATION DATA:
:       APPLICATION NUMBER: US/08/781,891
:       FILING DATE: 27-DEC-1996
:       CLASSIFICATION: 800
:       ATTORNEY/AGENT INFORMATION:
:         NAME: No. 6090620tenburg Ph.D., Carol
:         REGISTRATION NUMBER: 39,317
:         REFERENCE/DOCKET NUMBER: 240052.419
:         TELECOMMUNICATION INFORMATION:
:           TELEPHONE: (206) 622-4900
:           TELEFAX: (206) 682-6031
:         INFORMATION FOR SEQ ID NO: 207:
:           SEQUENCE CHARACTERISTICS:
:             LENGTH: 29604 base pairs
:             TYPE: nucleic acid
:             STRANDEDNESS: single
:             TOPOLOGY: linear
:   US-08-781-891-207

Query Match
Best Local Similarity 56.5%; Score 51.2; DB 3; Length 29604;
Matches 95; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 545 TTCTGTATCCCTCAAGTCTCCAACTCTTATTGAAGATTCAACACTTGTAAAGTAGG 604
DB 19789 TTATGTTTCCCCAGGAGATTAAATGTACTAGTAAGAAATCAATTAAGAGCAGG 19848
QY 605 TATTGAATGATGCTGCTCTGTGAAGCTTTTCCATGACTATGAGATTAGTATCAAGA 664
DB 19849 GGTGGGATGAGGGGACAGTGGAAACTTCTCGGATTTCACCTCAAGTTGGAGAG 19908
QY 665 TGTGAGACTTTTACAGATTTCAGCAACAAAATATGTCGAGATAA 712
DB 19909 TTTTGTGACACTGACGATCTTGCCTCAATGAAAAGTGGCGTAAATAA 19956

RESULT 7
US-08-232-463-14
: Sequence 14, Application US/08232463
: Patent No. 5670367
: GENERAL INFORMATION:
:   APPLICANT: DORNER, F.
:   APPLICANT: SCHEIFLINGER, F.
:   APPLICANT: FALKNER, F. G.
:   TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
:   NUMBER OF SEQUENCES: 52
```

```

: CORRESPONDENCE ADDRESS:
:   ADDRESSEE: Foley & Lardner
:   STREET: 1800 Diagonal Road, Suite 500
:   CITY: Alexandria
:   STATE: VA
:   COUNTRY: USA
:   ZIP: 22313-0299
:   COMPUTER READABLE FORM:
:     MEDIUM TYPE: Floppy disk
:     COMPUTER: IBM PC compatible
:     OPERATING SYSTEM: PC-DOS/MS-DOS
:     SOFTWARE: Patentin Release #1.0, Version #1.25
:     CURRENT APPLICATION DATA:
:       APPLICATION NUMBER: US/08/232,463
:       FILING DATE:
:       CLASSIFICATION: 435
:       PRIOR APPLICATION DATA:
:         APPLICATION NUMBER: US/07/935,313
:         FILING DATE:
:         APPLICATION NUMBER: EP 91 114 300.6
:         FILING DATE: 26-AUG-1991
:         ATTORNEY/AGENT INFORMATION:
:           NAME: BENT, Stephen A.
:           REGISTRATION NUMBER: 29,768
:           REFERENCE/DOCKET NUMBER: 30472/114 IMMU
:           TELECOMMUNICATION INFORMATION:
:             TELEPHONE: (703)836-9300
:             TELEFAX: (703)683-4109
:             TELEX: 899149
:           INFORMATION FOR SEQ ID NO: 14:
:             SEQUENCE CHARACTERISTICS:
:               LENGTH: 7218 base pairs
:               TYPE: nucleic acid
:               STRANDEDNESS: single
:               TOPOLOGY: linear
:             IMMEDIATE SOURCE:
:               CLONE: pTZ9pt-F15
:   US-08-232-463-14

Query Match
Best Local Similarity 4.7%; Score 49.2; DB 1; Length 7218;
Matches 33; Conservative 140; Mismatches 113; Indels 0; Gaps 0;

QY 9 ATTAATTTTATTTTGTTCAGTAAGAAATGTCATGTCMAATGATCGACGAC 68
DB 992 ACATATTTCTCGTTGTCATGCTCAGACAGATTAATTCGAGCTTGCTGACGTC 1051
QY 69 GCTTTACAGAGAGAGCTTCTGCTATGAGCGCCATCGAAGCTTCTACATTTCTCC 128
DB 1052 GAGGAGCTTCCGATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1111
QY 129 GGTCTCTCTCTCTCTCTCTCTCTGCTCGACGCTACAGCTACACCTCGTCAT 188
DB 1112 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1171
QY 189 GGCACGAGAGAGATCCAAATCAAAATCCCAATATATCCGTCGCCATGCTGTTCC 248
DB 1172 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1231
QY 249 ATCACTCTCTACATCTTATTAAGATTTCTCTCTCCCGTTGCC 294
DB 1232 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1277

RESULT 8
US-09-300-672-3
: Sequence 3, Application US/09300672
: Patent No. 6248937
: GENERAL INFORMATION:
:   APPLICANT: Finkelstein, Ruth R.
:   APPLICANT: Lynch, Tim
:   APPLICANT: Goodman, Howard M.
:   APPLICANT: Wang, Ming-Li
```



```
Sequence 1, Application US/08485216
Patent No. 5695960
GENERAL INFORMATION:
APPLICANT: Chan, Voon Loong
APPLICANT: Hanl, Eric K.
TITLE OF INVENTION: HIPURICASE GENE
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merchant & Gould
STREET: 3100 No. 5695960west Center, 90 S. 7th Street
CITY: Minneapolis
STATE: Minnesota
COUNTRY: U.S.A.
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,216
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 94/26907
FILING DATE: 13-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Mueller, Douglas P.
REGISTRATION NUMBER: 30,300
REFERENCE/DOCKET NUMBER: 7933.29US02
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612/332-5300
TELEFAX: 612/332-9081
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1338 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1338
US-08-485-216-1
```

```
Query Match 3.3%; Score 34.2; DB 1; Length 1338;
Best Local Similarity 47.8%; Pred. No. 0.67;
Matches 99; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 509 AGATAGTAATTATTTGATGATTTTTCATTTTCATTCCTGATCCCTCAAGTCCTCA 568
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 798 AGATCCATTTATGCTCTCTTCTTACTTATTTGTAACCTTACAAAGCATAGATCTGCAA 857
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 569 ACATCTTATGAGATTCACACCTGTAAGGTAGTATGATGATGATGATGATGATGATG 628
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 858 TGTGATCCCAAAATTCAGACAGTGTAGACATAGAGCTTTTATCTGACATGCTTT 917
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 629 GAAGCTTTCCATGACTAGTATGATGATGATGATGATGATGATGATGATGATGATG 688
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 918 TAATATCATTCACAGATTTTGAAGATTAAGTATGAGTATGAGATTAATATGAAC 977
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 689 CAACCAAAAAATTGGTGAGATAAAA 715
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 978 TAGAAAGCTAAGTGAAGAAAAAAA 1004
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
RESULT 12
US-09-003-245-1
; Sequence 1, Application US/09003245
; Patent No. 5981189
; GENERAL INFORMATION:
; APPLICANT: Chan, Voon L.
; APPLICANT: Hanl, Eric K.
```

```
TITLE OF INVENTION: Hipuricase Gene
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: MERCHANT & GOULD
STREET: 3100 No. 5981189west Center, 90 South Seventh Street
CITY: Minneapolis
STATE: Minnesota
COUNTRY: U.S.A.
ZIP: 55402-4131
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/003,245
FILING DATE: 06-JAN-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mueller, Douglas
REGISTRATION NUMBER: 30,300
REFERENCE/DOCKET NUMBER: M8G 7933.29-US-D2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (612) 332-5300
TELEFAX: (612) 332-9081
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1338 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Campylobacter jejuni
STRAIN: TGH9011
IMMEDIATE SOURCE: ATCC43431
LIBRARY: C. jejuni genomic library in pBR322
CLONE: PHIP-O
POSITION IN GENOME:
MAP POSITION: 380 kb SmaI fragment
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1338
US-09-003-245-1
```

```
Query Match 3.3%; Score 34.2; DB 2; Length 1338;
Best Local Similarity 47.8%; Pred. No. 0.67;
Matches 99; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 509 AGATAGTAATTATTTGATGATTTTTCATTTTCATTCCTGATCCCTCAAGTCCTCA 568
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 798 AGATCCATTTATGCTCTCTTCTTACTTATTTGTAACCTTACAAAGCATAGATCTGCAA 857
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 569 ACATCTTATGAGATTCACACCTGTAAGGTAGTATGATGATGATGATGATGATGATG 628
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 858 TGTGATCCCAAAATTCAGACAGTGTAGACATAGAGCTTTTATCTGACATGCTTT 917
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 629 GAAGCTTTCCATGACTAGTATGATGATGATGATGATGATGATGATGATGATGATG 688
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 918 TAATATCATTCACAGATTTTGAAGATTAAGTATGAGTATGAGATTAATATGAAC 977
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 689 CAACCAAAAAATTGGTGAGATAAAA 715
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 978 TAGAAAGCTAAGTGAAGAAAAAAA 1004
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
RESULT 13
US-08-853-552-1
; Sequence 1, Application US/08853552
; Patent No. 6013501
; GENERAL INFORMATION:
; APPLICANT: Chan, Voon Loong
```

```

; APPLICANT: Hani, Eric K.
; TITLE OF INVENTION: HIPURICASE GENE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant & Gould
; STREET: 3100 No. 601501west Center, 90 S. 7th Street
; CITY: Minneapolis
; STATE: Minnesota
; COUNTRY: U.S.A.
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/853,552
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/485,216
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mueller, Douglas P.
; REGISTRATION NUMBER: 30,300
; REFERENCE/DOCKET NUMBER: 7933.29US02
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612/332-5300
; TELEFAX: 612/332-9081
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1338 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1338
; US-08-853-552-1

Query Match
3.3%; Score 34.2; DB 3; Length 1338;
Best Local Similarity 47.8%; Pred. No. 0.67;
Matches 99; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 509 AGATGTAATTTATTTGTATGATATTTTCTGATCTGATCCCTCAAGTCTCCA 568
DB 798 AGATTCCTATTTATGCTGCTTCTTTCTTATTTGTAACATTAAGCATAGTATCTTGCAA 857
QY 569 ACATCTTATGTAAGATTCACACCTGTAAGGTAGTATTTGAATTTGATGACTCTGT 628
DB 858 TGTTCATATCCCAAAATTCAGCAGTTGTAAGCATAGGACTTTTAAATGCTGGACATGCTTT 917
QY 639 GAAGCTTTTCATGACTATGAGTAGTATCAAAAGATGTTGAGATCTTTTCAGATTTAGC 688
DB 918 TAAATATCATTCACATATTTGTAAGATTAATAATGAGTGTAGAGATTAGATAAAGAAGC 977
QY 689 CAACCAAAAATTTGTTGGAGATTAATA 715
DB 978 TAGAAAAGCTACTGAAGAAAAA 1004

RESULT 14
US-09-346-408-3/c
; Sequence 3, Application US/09346408B
; Patent No. 6338966
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Anderson, Shawn
; APPLICANT: Falcio, Carl
; APPLICANT: Rafalski, Antoni
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-A

```

```

; CURRENT APPLICATION NUMBER: US/09/346,408B
; CURRENT FILING DATE: 1999-07-01
; EARLIER APPLICATION NUMBER: 60/092,833
; EARLIER FILING DATE: July 14, 1998
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 3
; LENGTH: 2280
; TYPE: DNA
; ORGANISM: Momordica charantia
; US-09-346-408-3

```

```

Query Match
3.2%; Score 34; DB 4; Length 2280;
Best Local Similarity 63.4%; Pred. No. 1.1;
Matches 52; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 10 TTAATTTTATTTTGTGTTGTAAGAAATGTCATGTCATTAATTTGATGACGACG 69
DB 2268 TTTTATTTTATTTTATTTTATTTTAAATTAATACCTTCGTCATTAATGAGACATC 2209
QY 70 CTTTACAGAGAGAGCTCT 91
DB 2208 CCTTTCATCACAGAGATTCT 2187

```

```

RESULT 15
US-09-134-001C-762/c
; Sequence 762, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 762
; LENGTH: 1272
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
; US-09-134-001C-762

```

```

Query Match
3.2%; Score 33.8; DB 4; Length 1272;
Best Local Similarity 50.7%; Pred. No. 0.87;
Matches 111; Conservative 0; Mismatches 102; Indels 6; Gaps 1;

QY 503 AAGTGTAGTATTAATTTATGATGATATGATATTTTTCATTTCTGATATCCCTCAAG 562
DB 318 AAGTCAGAGTATTAATTTATGATATTTTCTTCTGACGATATTCACAGTATCACTAATAT 259
QY 563 TCTCCAACTTATTTGAAGATTCACACCTGTAAGTAG-----GATTTGAATTGA 616
DB 258 TCTCCAAATATTTTGAAGCCCTATATCTTAATAATGAATAATTAATTTGCAAAATGC 199
QY 617 TGTGACTCTGTGAAGCTTTTCATGACTAGAGATGATATCAAAAGATGATGAGATCT 676
DB 198 AAACAGCATTAATAATTTGTCAGTAAACACTGCTGATTTGACAGAAATGCTGAGACAT 139
QY 677 TTCAGATTTAGCCAAACCAAAATTTGTTGAGATTAATA 715
DB 138 AGCTGATGAGCATTAAATAAAACGTAACCATATATATA 100

```

Search completed: June 14, 2003, 08:24:33
Job time : 72 secs

Result	Score	Query	Match	Length	DB	ID	Description
No.							
1	11.1	10.6	5189	9	US-09-292-755-31		Sequence 31, Appl
2	11.1	10.6	5189	10	US-09-954-456-1131		Sequence 1131, Appl
3	44.4	4.2	3149	10	US-09-729-674-127		Sequence 127, Appl
4	41.4	3.9	941	9	US-10-123-155-464		Sequence 464, Appl
5	40.8	3.9	13059	7	US-08-781-986A-220		Sequence 220, Appl
6	38.2	3.6	439	10	US-09-864-761-20174		Sequence 20174, Appl
7	37	3.5	1910	9	US-09-938-842A-4570		Sequence 4570, Appl
8	36.2	3.5	5215	9	US-10-531-061-71		Sequence 71, Appl
9	36.2	3.5	6177	9	US-10-124-800-3		Sequence 3, Appl
10	36	3.4	1881	9	US-10-198-846-6503		Sequence 6503, Appl
11	36	3.4	99915	10	US-09-816-095-3		Sequence 3, Appl
12	36	3.4	659158	9	US-09-771-208-20		Sequence 20, Appl
13	35.8	3.4	29544	10	US-09-464-767-1		Sequence 1, Appl
14	35.8	3.4	32745	10	US-09-464-767-3		Sequence 3, Appl
15	35.4	3.4	127197	9	US-09-754-853A-1		Sequence 1, Appl
16	35	3.3	954	10	US-09-947-971-3		Sequence 3, Appl
17	34.6	3.3	280	9	US-09-933-797-413		Sequence 413, Appl
18	34.4	3.3	463	10	US-09-864-761-5350		Sequence 5350, Appl
19	34.4	3.3	2000	9	US-09-938-842A-4931		Sequence 4931, Appl

20	34.4	3.3	53326	10	US-09-818-664-3	Sequence 3, Appl1
21	34	3.2	5227	10	US-09-864-761-6599	Sequence 6599, App
22	34	3.2	527	10	US-09-864-761-8066	Sequence 8066, App
22	34	3.2	1827	10	US-09-985-502-270	Sequence 270, App
23	34	3.2	3769	9	US-10-175-523-184	Sequence 184, App
24	34	3.2	556	10	US-09-864-761-8619	Sequence 8619, App
25	33.8	3.2	734	9	US-10-184-644-458	Sequence 458, App
26	33.8	3.2	734	9	US-10-184-634-458	Sequence 458, App
27	33.8	3.2	785	10	US-09-770-445-850	Sequence 450, App
28	33.8	3.2	98829	9	US-10-017-724-3	Sequence 3, Appl1
29	33.8	3.2	288	10	US-09-878-574-10134	Sequence 10134, App
30	33.6	3.2	125	10	US-09-878-574-7317	Sequence 7317, App
31	33.6	3.2	833	9	US-08-781-9864-4473	Sequence 4473, App
32	33.6	3.2	5814	10	US-10-092-1514-1860	Sequence 1860, App
33	33.6	3.2	5814	10	US-09-764-847-1860	Sequence 1860, App
34	33.6	3.2	533	10	US-09-864-761-6759	Sequence 6759, App
35	33.4	3.2	556	9	US-09-918-995-27130	Sequence 27130, App
36	33.4	3.2	2979	10	US-09-815-242-4754	Sequence 4754, App
37	33.4	3.2	3009	10	US-09-815-242-8801	Sequence 8801, App
38	33.4	3.2	7035	10	US-09-815-242-9060	Sequence 9060, App
39	33.4	3.2	277	9	US-09-815-242-8615	Sequence 8615, App
40	33.2	3.2	277	9	US-09-535-459-1301	Sequence 1301, App
41	33.2	3.2	477	9	US-09-864-761-5436	Sequence 5436, App
42	33.2	3.2	225883	9	US-10-175-523-57	Sequence 57, Appl
43	33.2	3.2	305	10	US-09-864-761-19262	Sequence 19262, App
44	33	3.1	496	10	US-09-864-761-2534	Sequence 2534, App
45	33	3.1				

ALIGNMENTS

RESULT 1
US-09-292-758-31

; Sequence 31, Application US/0929292758
; Publication No. US20020197602A1

```

; GENERAL INFORMATION:
; APPLICANT: Burmer, Glenna C.
; ADDITIONAL: Burmer, Glenna C.

```

;; APPLICANT: Brown, Joseph P.
;; APPLICANT: Lifespan Biosciences, Inc.
;; TITLE OF INVENTION: Nucleic Acid Sequencing

FILE REFERENCE: 017473-001110US

;; CURRENT APPLICATION NUMBER: US/09/2
;; CURRENT FILING DATE: 1999-04-14

; EARLIER APPLICATION NUMBER: US 60/C
; EARLIER FILING DATE: 1998-04-15

```

; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: FastSeq for Windows Version 1.0.0

```

```

; SEQ ID NO 31
; LENGTH: 5189
; TYPE: DNA

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
ITS-09-292-758-31

```

Query Match 10.6%; Score

Best Local Similarity 54.1%; Prec
Matches 272; Conservative 0; M

QY 396 GATACCAAGAGAGATGAATCTGGP

Db 433 GATATTAGCATGAGTCTATCAGAT

456 AGTTT⁺TAGAAAAGGTGTCTCTCCG

Db 493 TTATACATAGAGGAACT---T

QY 516 AATTATTGTGATGTTATGCATATT
| | | | | | | | | |

Db 550 AGCAATGTTACTTGTCCACGTT

QY 573 CTTATGAAGATTCAACACTTGT
| | | | | | | | | |

DD 11GCTTGAATAATATAGCTAGT1TAA

610 TTGCTTGAATAATAGCAGTTATTTAAAGGCAGGTGTAGGATTTGAAGGAGATCAGTGGAA 665

QY	63	CTTTCCATGACATGAGACTTGGTATGCCAAATGTTGAGATCTTTAGATTAGCCAAAC	632
	670	CTTCTACGTGACTTGTGATATTCAAATTTGAAGATTTTGTGGAGTTGCACAGATGTGCCAAT	729
QY	693	CAAAAAATGTGTGAGATAAAAAATGTGGGGCCCTTGCCTCATCACTGAGACACTGTTGCG	752
	730	AAAAAGCTGAAATGTACACAGACCTGTGAGCCCTTACAGCTCTGTGTTAAACACCTCTTAGGT	789
QY	753	AAAGACTCTCTGAAGCCAAACAGAAATCAAGCCTTGGAACTGGGAGTTTATCCCTGTCA	812
	790	AAACACCTCTCTGAAGACAAAGTCTATCCGCTGTACCAATTTGAGATTAATTTCCCTCTACT	849
QY	813	AACGACAGTTAAATCAATACGACAGCAACGAGTCTTATGCTTCATGGCACTCTTACAAAGTT	872
	850	GAGGACCAAGAAACTGTATGCACAGCCACTGATGTTATGCTGTTTATATATTACCGAAT	909
QY	873	CTTAAGGACCTTCTGATGCTGT	895
	910	TTAGAGATTTTGGATGATGATCTGT	932

```

RESULT 2
US-09-954-456-1131
: Sequence 1131, Application US/09954456
: Patent No. US20020115057A1
: GENERAL INFORMATION:
: APPLICANT: Young, Paul
: TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Can
: TITLE OF INVENTION: Sets
: FILE REFERENCE: 689290-76
: CURRENT APPLICATION NUMBER: US/09/954,456
: CURRENT FILING DATE: 2001-09-18
: PRIOR APPLICATION NUMBER: US/60/233,617
: PRIOR FILING DATE: 2000-09-18
: PRIOR APPLICATION NUMBER: US/60/234,052
: PRIOR FILING DATE: 2000-09-20
: PRIOR APPLICATION NUMBER: US/60/234,923
: PRIOR FILING DATE: 2000-09-25
: PRIOR APPLICATION NUMBER: US/60/235,134
: PRIOR FILING DATE: 2000-09-25
: PRIOR APPLICATION NUMBER: US/60/235,637
: PRIOR FILING DATE: 2000-09-26
: PRIOR APPLICATION NUMBER: US/60/235,638
: PRIOR FILING DATE: 2000-09-26
: PRIOR APPLICATION NUMBER: US/60/235,711
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: US/60/235,720
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: US/60/235,840
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: US/60/235,863
: PRIOR FILING DATE: 2000-09-27
: NUMBER OF SEQ ID NOS: 2276
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 1131
: LENGTH: 5189
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-954-456-1131

```

Query Match	10.6%	Score 11;	DB 10;	Length 5189;
Best Local Similarity	54.1%;	Pred. No. 2e-21;		
Matches 272;	Conservative 0;	Mismatches 225;	Indels 6;	Gaps 2;
QY	396	GATACCAAGAGAGATGAACTCTGGAACTTTTGCTGGCTGGATATTGAGTGGAGACCA	455	
DB	433	GATATTAGCAAGACTATCAGATGGGGATGTGGGGATTGGAGATTGACATGAGAGGCCACCA	492	
QY	456	AGTTTGGAAAAGGTGTTCTCCCGGGGAAGTTGGCACTGCCAGATATGTGTAGATAGT	515	
DB	493	TTTATACATACAGGGAACCT--TGGCAGAACTGTGCATTAATCTAGTTGTGGTTCTGAG	549	

OY	510	AATTATTGGAGCTTTTGATATTTTT---CATCTGATACCCCTCAAACTCCCAACAT	512
Db	550	AGCAAAATGGTACTTGTTCCAGCTTTCCCATGATGATTTTTTCCCGAGGATTTAAAAATG	609
OY	573	CTTATTGAAGATTCAACACTTTAAAGTAGTATTGAAATGATAGTCACTGCTGTAAG	632
Db	610	TTGCTTGAAATTAAGCACTTTAAAAAGCAGGTGTAGAAATTGAAGAGATCACTGTAAG	669
OY	633	CTTTTCCAGTACTAGGAGTTAGTATCAAAAGATGTTAGAGATCTTTCAGATTTTAGCCAAC	652
Db	670	CTCTACGTGACTTTGATATCAAAATTTGAAGAAATTTTGGAGTTGACAGATGTTGCCAAT	729
OY	693	CAAAAAATTTGGTGGAGATAAAAAATGGGGCCCTTGGCTACATACTGAGACACTGTTTC	752
Db	730	AAAAAAGCTGAATGTACAGAGACCTGGAGCCCTTAACACTCGTTTAAACACTCTTAAGT	789
OY	753	AAAGAGCTCCTGAAGCCAACAGAAATCAAGGCTTGGGAAGCTGGAGTTTATTCCTGTCA	812
Db	790	AAACAGCTCTGAAGACACAAGTCTTATCCGCTGTAGCAATTTGGAAATTAATTTCTCTCACT	849
OY	813	AAGCAGCACTTACAAATACAGACCAACGGATGTAATGCTTCANTGGCATCTTACAAAGTT	872
Db	850	GAGGACCAAGAACTGATGTACAGCCCACTGATGCTTATTTATTAATTAACCAAAAT	909
OY	873	CTTAAGGACCTTCCATGCTCTG	895
Db	910	TTAGAGATTTTGGATGATGACTGT	932

RESULT 3
 US-09-729-674-127
 ? Sequence 127: Application US/09729674
 ? Patent No. US20010039335A1
 ? GENERAL INFORMATION:
 ? APPLICANT: Jacobs, Kenneth
 ? APPLICANT: McCoy, John M.
 ? APPLICANT: Lavallee, Edward R.
 ? APPLICANT: Collins-Racie, Lisa A.
 ? APPLICANT: Evans, Cheryl
 ? APPLICANT: Merberg, David
 ? APPLICANT: Treacy, Maurice
 ? APPLICANT: Agostino, Michael J.
 ? APPLICANT: Steinger II, Robert J.
 ? APPLICANT: Spaulding, Vikki
 ? APPLICANT: Wong, Gordon G.
 ? APPLICANT: Clark, Hilary
 ? APPLICANT: Fechtel, Kim
 ? APPLICANT: Genetics Institute, Inc.
 ? TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
 ? FILE REFERENCE: 6055-64X
 ? CURRENT APPLICATION NUMBER: US/09/729,674
 ? CURRENT FILING DATE: 2000-12-04
 ? PRIOR APPLICATION NUMBER: 09/539,330
 ? PRIOR FILING DATE: 2000-03-30
 ? NUMBER OF SEQ ID NOS: 283
 ? SOFTWARE: patentIn Ver. 2.0
 ? SEQ ID NO 127
 ? LENGTH: 3149
 ? TYPE: DNA
 ? ORGANISM: Homo sapiens
 US-09-729-674-127

	Query Match	4.2%	Score 44.4	DB 10	Length 3149
	Best Local Similarity	49.3%	Pred. No. 0.045		
	Matches 148	Conservative	0	Mismatches 146	Indels 6
					Gaps 1
Qy	571 ATCTTATTGAAGATTCAACACTTGTAAAGGTAGTATTGGCAATTATGCTGACTCTGTGCA				630
Db	470 ATATTTTGGCAGATGGCCACCAATTTTGAAGTTGGAGTGGAGTCTCAAGAGTGGCAGCA				529
Qy	631 AGCTTTTCCAGACTATGAGTATGATCAAGAAGTGTGAGAGATCTTTCAGATTATTCGCA				690
Db	530 AGCTTTTCAGAGATTATATGCTCGTGTGTAGGGGGTGCCTGGACCTCGATATCTCAGCA				589

QY 691 -----ACCAAAATTTGGTGGAGATAAATGGGGCTTGCCTCACTAAGTACAGACAC 744
 DB 590 TCGGCGAGAGAAACATTTGCTCTGTAAAGGCGTTAGCCTGAAGTCCCTCGCTGAGACG 649
 QY 745 TTGTTTGAAGAGCTCTGAAAGCCAAAGATCAGGCTTGGGAAGTGGATTTATC 804
 DB 650 TTTTGAACCTTTCCCTTGAAGTCCCTTCTACTTCTTGACAGCAACTGGATGTGAGA 709
 QY 805 CTCTGTCAAGAGAGATTCATATAGCGACAGAGATGCTTATGCTTATGCAATCTT 864
 DB 710 CTCTCAGAGAGAGAGATTTATTTATGCTGCGAGGATGCGAGATTCAGTGGCTCTCT 769

RESULT 4
 US-10-123-155-464/C
 ; Sequence 464, Application US/10123155
 ; Publication No. US20030068794A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: Deforge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvarioff, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerlitsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Guiney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Matanabe, Colin K
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3330R1C30
 ; CURRENT APPLICATION NUMBER: US/10/123,155
 ; PRIOR APPLICATION removed - See Palm or File Wrapper
 ; NUMBER OF SEQ ID NOS: 550
 ; SEQ ID NO 464
 ; LENGTH: 941
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 ; US-10-123-155-464

Query Match 3.9%; Score 41.4; DB 9; Length 941;
 Best Local Similarity 7.7%; Pred. No. 0.15; Mismatches 476; Indels 0; Gaps 0;
 Matches 58; Conservative 215; Mismatches 476; Indels 0; Gaps 0;
 QY 68 CCGTTTACAGAGAGAGCTTCTGCTATCGACGCGCAAGCTTCTTACATTTCTC 127
 DB 870 SSSC...M.B.MNMNY...WT..RCB.NYC.A.....AM.HCM.S.H..W..M.M.811
 QY 128 CCGTCTTTTCTTCTTCTTCTGCTGCTCGACCGTACAGCTCAACCTCCGTCCA 187
 DB 810 N.AYG.T...SM.ASS.S..RMSR..HMC.AS.TCOT.BT.ABHB..S.NCMS.MMY.751
 QY 188 TGGCAGAGAGAGAGCTTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 247
 DB 750 RC.TY.BG..RNDG...ST.K..SBSC.HAWA.MH..HY..Y..TM..A.B.NKH.691
 QY 248 CATCACTTCTTACATCTTAAAGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 307
 DB 690 YM.K.MRK...N.C..B.K..A.DM.R.S.H.TM..S.MC.SB...TNN..STYNSB.631
 QY 308 TCCAGCATGAGGTTTGGTGGTGAATTTTGTATACCAAGACTGCTAGAGTTGATTA 367
 DB 630 TADACM..CA.SHWCH.RDB.RCNKCN.M.W.B.....BHAAAM...YDBKMSMA.571

QY 368 GCGAGCATGACCTTATTAAGTTCTGTATACCAAGAGATGATTCGATAGCTTT 427
 DB 570 ..A..BDW.RCH.FCHSCKRD..MKDNYCBA.A...CY..AMAKKMBHC..D.511
 QY 428 TGTGCTTGTATATTGAGTGGAGACCAAGTTTGAAGAGTGTCTCCGGGAGGT 487
 DB 510 WSSSSSD..SYSG.CHKMBCHA.G..STKSHW.H.NMAMRSDM..R.B.CSM.THTS.R.451
 QY 488 TGGGACTGTCAGATATGAGTATGATATTTGTGATGATGATGATGATGATGATGATGATG 547
 DB 450 .Y.KN..GTOMHRSBH..K.Y..T.N.B.ASB.DSSN.THB.KTH.GMC..RCHB...D.391
 QY 548 TGGTATCCCTCAAGAGCTTCAACATCTTATTTGAAGATTCACACTGTAAGTAGTAT 607
 DB 390 ABSBSB..K.MT.C.N.W..HNW.KAB.NC.W.DT..DTBAB.C.MSSTSM.TH...T.331
 QY 608 TGGATTTGATGAGTACTGTGTGAGGCTTTCCATGACTATGAGTATGATCAAGATGT 667
 DB 330 S.YRAA.CWN.KTCS..H..TT.H.M...R..S.RH.R...ABTH.TRHT.N.MH.BT.271
 QY 668 TGAAGATCTTTCAGATTAGCCACCAAAATTTGGTGGAGATTAATTTGGGCTTGC 727
 DB 270 RBSBMCMA.MBS..HS...TB.RASKBABH.DH...C.TBBSMB..KNS.T.DY.211
 QY 728 CTCACTAAGTACAGACCTTGTTCGAAAGAGCTCTGAAAGCCAAAGATCAGGCTTG 787
 DB 210 .YX.M.S.STM.T.H.G..TKYTTA...AST..Y..C.MATRASMR.CD.A.S.NCR.151
 QY 788 GAAGTGGAGTTTATCTCTGTCAAGC 816
 DB 150 D.BBAR..CB...T..T.....D..B.122

RESULT 5
 US-08-781-986A-220
 ; Sequence 220, Application US/08781986A
 ; Publication No. US20030054436A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Charles Kunsch
 ; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
 ; NUMBER OF SEQUENCES: 5255
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville
 ; STATE: Maryland
 ; COUNTRY: USA
 ; ZIP: 20850
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
 ; COMPUTER: HP Vectra 486/33
 ; OPERATING SYSTEM: MSDOS version 6.2
 ; SOFTWARE: ASCII Text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/781,986A
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Benson, Bob
 ; REGISTRATION NUMBER: 30,446
 ; REFERENCE/DOCKET NUMBER: PB248PP
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (301) 309-8504
 ; TELEFAX: (301) 309-8512
 ; INFORMATION FOR SEQ ID NO: 220:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 13059 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear


```

; LOCATION: 1753, 1774, 1785, 1792, 1801, 1808, 1818, 1823, 1827, 1829
; LOCATION: 1834, 1843, 1856, 1867

```

Dd 390959 CTTCTCCTCTTCCTCTTCTCTTCCTN

Oy 232 GCCAATGGCCCTGTTCCATCACTTCTACATCTTAATAAGCATTTCCCTCCTCC 287
||| ||| | ||| ||| ||| ||| ||| |||
Db 391019 TTCACTACTTCATATTACATGTACGCTCCACCTCTTAACAATAATCTCTCCCTCC 391074

RESULT 13
US-09-464-767-1

```

: Sequence 1, Application US/09464767
: Patent No. US20020045249A1
: GENERAL INFORMATION:
: APPLICANT: Both, Gerald
: APPLICANT: Boyle, David
: APPLICANT: Virati, Sudhanshu
: TITLE OF INVENTION: DNA Encoding Ovine Adenovirus (OAV287) and Its Use as a Viral Vec
: FILE REFERENCE: 50179-073
: CURRENT APPLICATION NUMBER: US/09/464,767
: CURRENT FILING DATE: 1999-12-16
: NUMBER OF SEQ ID NOS: 3
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 1
: LENGTH: 29544
: TYPE: DNA
: ORGANISM: Ovine adenovirus
: US-09-464-767-1

```

Query Match	3.4%	Score 35.8	DB 10	Length 29544
Best Local Similarity	50.9%	Pred. No. 65		
Matches 85; Conservative	0	Mismatches 82	Indels 0	Gaps 0

QY 442 TTGAGTGGAGACCAAGTTT TAGAAAAGTGTTCTCTCCGGGGAAGTTGCGACTGTCCAGA 501

Db 6587 TTGATTCAGAGACCATGATTTTCACAAACGTTGACCAGTATTTTTTTACCTACTGAAATTT 6646

502 TATGCTAGATAATTATTGTGATGTTATGCATATTTTTCATTCTGCGTATCCCTCAA 561

Db 6647 GTAGCTCTGTTCTGAATTTTAAAGCTATGCCAATTCATCTTCTTTTATTTAACA 6706

Db 6707 TTACAAACATTCCTCTGTTTACCTCATACCTATATCGGTAGCTATT 6753

RESULT 14

```

US-09-464-767-3
Sequence 3, Application US/09464767
Patent No. US20020045249A1
GENERAL INFORMATION:
APPLICANT: Boyle, Gerald
APPLICANT: Boyle, David
APPLICANT: Vratil, Sudhanshu
TITLE OR INVENTION: DNA Encoding Ovine Adenovirus (OAV287) and Its Use as a Viral Vec
FILE REFERENCE: 50179-073
CURRENT APPLICATION NUMBER: US/09/464,767
CURRENT FILING DATE: 1999-12-16
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn version 3.0
SEQ ID NO: 3
LENGTH: 32745
TYPE: DNA
ORGANISM: synthetic construct
US-09-464-767-3

```

Query Match	3.4%	Score 35.8	DB 10	Length 32745
Best Local Similarity	50.9%	Pred. No. 70		
Matches 85; Conservative	0	Mismatches 82	Indels 0	Gaps 0

442 TTGAGTGGAGACCAAGTTT TAGAAAGGTGTCTCCCGGGAAGTTGCGACTGTCCAGA 501

Db 6586 TTGATTCAAGACCATTGATTTTCACAAACGTTGACCAGTATTTTCTAGCTACTGAATTT 6645

502 TATGTGATAGTAATTATTTGTGATGTTATGCATATTTTTCATTTCTGGTATCCCTCAA 561

Db 6646 GTAGCTCTGTCTGAATTTTAAAGCTATGCCAATTCATCTCTTTTATTAAACA 6705

QY	562	GTC	CCA	ACAT	CTT	ATG	AAG	ATC	CAAC	ACT	TG	TAA	AGG	TAG	GAT	ATT	608
Db	6706	TTA	CAAA	ACAT	CTCT	GT	TAC	CTC	ATA	ACCT	TAT	ATC	CG	TAG	GCT	ATT	6752

RESULT 15
US-09-754-853A-1

```

? Sequence 1, Application US/09754853A
? Publication No. US20030005451A1
?
? GENERAL INFORMATION:
?
? APPLICANT: Hauge, Brian M.
? APPLICANT: Parnell, Laurence D.
? APPLICANT: Parsons, Jeremy D.
? APPLICANT: Wang, Ming Li
? TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
? TITLE OF INVENTION: Soybean Cyst Nematode Resistance
? FILE REFERENCE: 38-10(15810)B
? CURRENT APPLICATION NUMBER: US/09/754,853A
? CURRENT FILING DATE: 2001-01-05
? PRIOR APPLICATION NUMBER: US 60/174,880
? PRIOR FILING DATE: 2000-01-07
? NUMBER OF SEQ ID NOS: 1119
?
? SEQ ID NO 1
? LENGTH: 127197
?
? TYPE: DNA
?
? ORGANISM: Glycine max
?
? FEATURE:
?
? OTHER INFORMATION: Clone ID: 515002_region_G2
US-09-754-853A-1

```

Query Match	3.4%;	Score 35.4;	DB 9;	Length 127197;
Best Local Similarity	50.9%;	Pred. No. 2.2e+02;		
Matches 84;	Conservative 0;	Mismatches 81;	Indels 0;	Gaps 0;

501 ATATGCTAGATAGTAATTATTTGATGTTATGCATATTTTTCATTCTGGTATCCCTCAA 560

Db 18044 ATAGTTAAATAGATAAATGTTTGGATTGTTTGGATATTTTTTTAACAATAATCGTTGAA 1810

561 AGTCTCCAACATCTTATTGAAGATTCAACACTTGTAAAGGTAGGTAATTGGAATTGATGGT 620

Db 18104 TATCAAAACTTATTATTATTTTATAGGACGAGGAAGTCTGTAATGAAAGGAAGT 1816

621 GACTCTGTGAAGCTTTCCATGACTATGGAGTTAGTATCAAGAT 665

Db 18164 AATAAATTGAAGTTTATCAAAACCAATGCTTTCGGCTTCGAAAT 18208

Search completed: June 14, 2003, 09:31:29
Job time : 225 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2003, 07:03:29 ; Search time 2836 Seconds
(without alignments)
9299.883 Million cell updates/sec

Title: US-09-896-186b-23
Perfect score: 1049
Sequence: 1 accaagaacttaattttat.....tttaaccgcgcgaactag 1049

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 24791104 seqs, 12571243825 residues

Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending_Patents_NA_Main:*

- 1: /cgn2_6/ptodata/1/pna/US06.COMB.seq.*
- 2: /cgn2_6/ptodata/1/pna/US06.COMB.seq.*
- 3: /cgn2_6/ptodata/1/pna/US07.COMB.seq.*
- 4: /cgn2_6/ptodata/1/pna/US08.COMB.seq.*
- 5: /cgn2_6/ptodata/1/pna/US081.COMB.seq.*
- 6: /cgn2_6/ptodata/1/pna/US082.COMB.seq.*
- 7: /cgn2_6/ptodata/1/pna/US083.COMB.seq.*
- 8: /cgn2_6/ptodata/1/pna/US084.COMB.seq.*
- 9: /cgn2_6/ptodata/1/pna/US085.COMB.seq.*
- 10: /cgn2_6/ptodata/1/pna/US086.COMB.seq.*
- 11: /cgn2_6/ptodata/1/pna/US087.COMB.seq.*
- 12: /cgn2_6/ptodata/1/pna/US088.COMB.seq.*
- 13: /cgn2_6/ptodata/1/pna/US089.COMB.seq.*
- 14: /cgn2_6/ptodata/1/pna/US090.COMB.seq.*
- 15: /cgn2_6/ptodata/1/pna/US091.COMB.seq.*
- 16: /cgn2_6/ptodata/1/pna/US092.COMB.seq.*
- 17: /cgn2_6/ptodata/1/pna/US093.COMB.seq.*
- 18: /cgn2_6/ptodata/1/pna/US094.COMB.seq.*
- 19: /cgn2_6/ptodata/1/pna/US095A.COMB.seq.*
- 20: /cgn2_6/ptodata/1/pna/US095B.COMB.seq.*
- 21: /cgn2_6/ptodata/1/pna/US095C.COMB.seq.*
- 22: /cgn2_6/ptodata/1/pna/US095D.COMB.seq.*
- 23: /cgn2_6/ptodata/1/pna/US096A.COMB.seq.*
- 24: /cgn2_6/ptodata/1/pna/US096B.COMB.seq.*
- 25: /cgn2_6/ptodata/1/pna/US096C.COMB.seq.*
- 26: /cgn2_6/ptodata/1/pna/US096D.COMB.seq.*
- 27: /cgn2_6/ptodata/1/pna/US096E.COMB.seq.*
- 28: /cgn2_6/ptodata/1/pna/US097A.COMB.seq.*
- 29: /cgn2_6/ptodata/1/pna/US097B.COMB.seq.*
- 30: /cgn2_6/ptodata/1/pna/US097C.COMB.seq.*
- 31: /cgn2_6/ptodata/1/pna/US098A.COMB.seq.*
- 32: /cgn2_6/ptodata/1/pna/US098B.COMB.seq.*
- 33: /cgn2_6/ptodata/1/pna/US098C.COMB.seq.*
- 34: /cgn2_6/ptodata/1/pna/US099A.COMB.seq.*
- 35: /cgn2_6/ptodata/1/pna/US099B.COMB.seq.*
- 36: /cgn2_6/ptodata/1/pna/US099C.COMB.seq.*
- 37: /cgn2_6/ptodata/1/pna/US099D.COMB.seq.*
- 38: /cgn2_6/ptodata/1/pna/US100A.COMB.seq.*
- 39: /cgn2_6/ptodata/1/pna/US100B.COMB.seq.*
- 40: /cgn2_6/ptodata/1/pna/US101A.COMB.seq.*
- 41: /cgn2_6/ptodata/1/pna/US101B.COMB.seq.*
- 42: /cgn2_6/ptodata/1/pna/US102A.COMB.seq.*
- 43: /cgn2_6/ptodata/1/pna/US102B.COMB.seq.*

Result				SUMMARIES			
No.	Score	Query Match	Length DB	ID	Description		
1	1049	100.0	1049	33	US-09-896-186b-23	Sequence 23, Appl	
2	1039	99.0	1114	19	US-09-513-996A-67235	Sequence 67235, A	
3	1002.2	95.5	1129	19	US-09-513-996A-28855	Sequence 28855, A	
4	831	79.2	942	28	US-09-708-427-27137	Sequence 27137, A	
5	831	79.2	942	33	US-09-896-186b-1	Sequence 1, Appl1	
6	471.6	45.0	507	24	US-09-620-393B-6112	Sequence 6112, Ap	
7	302.2	28.8	118718	20	US-09-534-859-579	Sequence 579, Ap	
8	302.2	26.3	1170	34	US-09-803-736-579	Sequence 736, Ap	
9	276.2	26.3	1170	34	US-09-803-736-579	Sequence 736, Ap	
10	276.2	26.3	1170	65	US-60-218-993-15	Sequence 15, Appl	
11	181.2	17.3	672	23	US-09-605-698-9441	Sequence 9441, Ap	
12	174.8	16.7	584	63	US-09-849-529A-2068	Sequence 2068, Ap	
13	174.8	16.6	1434	28	US-09-708-427-65909	Sequence 65909, A	
14	174	16.6	1434	28	US-09-708-427-83271	Sequence 83271, A	
15	174	16.6	1434	28	US-09-803-736-1476	Sequence 1476, Ap	
16	151.2	14.4	130299	17	US-09-304-517A-210624	Sequence 210624, Ap	
17	142.4	13.6	579	17	US-09-371-146A-210624	Sequence 210624, Ap	
18	142.4	13.6	579	37	US-09-985-678-210624	Sequence 210624, Ap	
19	142.4	13.6	579	56	US-60-125-818-10555	Sequence 10555, A	
20	142.4	13.6	579	56	US-60-125-818-10555	Sequence 10555, A	
21	142	13.5	82484	20	US-09-534-859-326	Sequence 326, Appl	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

```
22 142 13.5 82484 31 US-09-803-736-326 Sequence 326, App
23 124.6 11.9 278 16 US-09-263-191-14639 Sequence 14639, A
24 124.6 11.9 278 17 US-09-304-517A-155805 Sequence 155805,
25 124.6 11.9 278 16 US-09-371-146A-155805 Sequence 155805,
26 124.6 11.9 278 37 US-09-975-254-14639 Sequence 14639, A
27 124.6 11.9 278 37 US-09-985-678-155805 Sequence 155805,
28 124.6 11.9 432 25 US-09-654-617-30344 Sequence 30344, A
29 124.6 11.9 432 25 US-09-684-016-16008 Sequence 16008, A
30 113.6 10.8 537 25 US-09-684-016-16008 Sequence 16008, A
31 113.6 10.8 537 25 US-09-684-016-16008 Sequence 16008, A
32 112.2 10.7 697 31 US-09-505-532-9796 Sequence 9796, Ap
33 112.2 10.7 697 31 US-09-819-091A-9796 Sequence 9796, Ap
34 111 10.6 4299 33 US-09-896-186b-17 Sequence 17, Appl
35 111 10.6 5189 71 PCT-US99-08314-31 Sequence 8683, Ap
36 111 10.6 5189 16 US-09-292-758-31 Sequence 31, Appl
37 111 10.6 5189 18 US-09-442-589B-1176 Sequence 1176, Ap
38 111 10.6 5189 61 US-09-954-456-1131 Sequence 1131, Ap
39 111 10.6 5189 61 US-60-172-373-7812 Sequence 7812, Ap
40 111 10.6 5189 65 US-60-213-360-4319 Sequence 4319, Ap
41 111 10.6 5189 76 PCT-US02-03574-11 Sequence 27164, A
42 111 10.6 5208 1 US-08-594-242-70 Sequence 70, Appl
43 111 10.6 5208 9 US-09-618-166-70 Sequence 70, Appl
44 111 10.6 5208 23 US-09-618-166-70 Sequence 70, Appl
45 111 10.6 5208 23 US-09-618-166-70 Sequence 70, Appl
```

ALIGNMENTS

```
RESULT 1
US-09-896-186b-23
; Sequence 23, Application US/09896186B
; GENERAL INFORMATION:
; APPLICANT: Joshua Z. Levin
; APPLICANT: Ken Phillips
; APPLICANT: Greg Budziszewski
; APPLICANT: Fred Meins
; APPLICANT: Zhenya Glazov
; TITLE OF INVENTION: Methods of Controlling Gene Expression
; FILE REFERENCE: PB/5-31481A
; CURRENT APPLICATION NUMBER: US/09/896,186B
; CURRENT FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 1049
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-896-186b-23
```

```
Query Match 100.0%; Score 1049; DB 33; Length 1049;
Best Local Similarity 100.0%; Pred. No. 5,8e+286;
Matches 1049; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ACCAAGCATTAATTTTATTTTGTTCAGTAAAGAAATGTCATCGTAAATTGGA 60
DB 1 ACCAAGCATTAATTTTATTTTGTTCAGTAAAGAAATGTCATCGTAAATTGGA 60
QY 61 TCGAGAGCGCTTTTACAGAGAGAGCTTCGATGAGCGCATCGAAGCTTCTTACA 120
DB 61 TCGAGAGAGCGCTTTTACAGAGAGAGCTTCGATGAGCGCATCGAAGCTTCTTACA 120
QY 121 ATTTCGCCGTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
DB 121 ATTTCGCCGTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
QY 181 CCGTCCATGCGCAGAGAGAGATCCAAATCCCAATATATTCGTCGCCAATTGC 240
DB 181 CCGTCCATGCGCAGAGAGAGATCCAAATCCCAATATATTCGTCGCCAATTGC 240
QY 241 CTCGTCATGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
DB 241 CTCGTCATGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
```

```
QY 301 GGAATTTTCCACCAATGAGCTTGTGTGAGATTTTGTATAGCAAGACTGCTACTGAGC 360
DB 301 GGAATTTTCCACCAATGAGCTTGTGTGAGATTTTGTATAGCAAGACTGCTACTGAGC 360
QY 361 TTGATTAAGCGACCAATGACGCTTATTAAGCTTGTATACCAAGAGATGATCTGAA 420
DB 361 TTGATTAAGCGACCAATGACGCTTATTAAGCTTGTATACCAAGAGATGATCTGAA 420
QY 421 TGCCTTTTGTGCTTGTATGATGATGAGAGCAAGTTTGAAGAGGTGTTCCCGCG 480
DB 421 TGCCTTTTGTGCTTGTATGATGATGAGAGCAAGTTTGAAGAGGTGTTCCCGCG 480
QY 481 GGAAGTTCCGACTGTCAGATATGTAGATAGTATATTTGTATGATGATATTT 540
DB 481 GGAAGTTCCGACTGTCAGATATGTAGATAGTATATTTGTATGATGATATTT 540
QY 541 TTGATTTCTGATTCCTCTCAAGCTTCAACATCTTATTAAGATTCACACTTGTAAAG 600
DB 541 TTGATTTCTGATTCCTCTCAAGCTTCAACATCTTATTAAGATTCACACTTGTAAAG 600
QY 601 TAGGATTTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 601 TAGGATTTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 661 AAGATGTTGAGATCTTTTCAGATTTAGCCACCAAAAAATGTTGAGATAAAAATGGG 720
DB 661 AAGATGTTGAGATCTTTTCAGATTTAGCCACCAAAAAATGTTGAGATAAAAATGGG 720
QY 721 GCCTTGCTCCTCACTAATGAGACACTGTTTGCAGAGAGCTCCTGAAGCCAAACAGATTA 780
DB 721 GCCTTGCTCCTCACTAATGAGACACTGTTTGCAGAGAGCTCCTGAAGCCAAACAGATTA 780
QY 781 GGCTTGAGGAGAGTGGAGTTTATCTCTCTCTCAAGAGAGAGTATACATAGCGAGACAGG 840
DB 781 GGCTTGAGGAGAGTGGAGTTTATCTCTCTCTCAAGAGAGAGTATACATAGCGAGACAGG 840
QY 841 ATGCTTATGCTTATGAGCATCTTTCACAAAGCTTCTTAAGAGACCTTCTGATGCTCAGTG 900
DB 841 ATGCTTATGCTTATGAGCATCTTTCACAAAGCTTCTTAAGAGACCTTCTGATGCTCAGTG 900
QY 901 GCTCATACGTAAGAGAGAGAGAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAA 960
DB 901 GCTCATACGTAAGAGAGAGAGAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAA 960
QY 961 ATGATATGATACACCAATCTATGATCAAGTATGATGATGATGATGATGATGATGAT 1020
DB 961 ATGATATGATACACCAATCTATGATCAAGTATGATGATGATGATGATGATGATGAT 1020
QY 1021 TCTGTCCTCTTAACCGTCCAGAAACTAG 1049
DB 1021 TCTGTCCTCTTAACCGTCCAGAAACTAG 1049
```

```
RESULT 2
US-09-513-996A-67235
; Sequence 67235, Application US/09513996A
; GENERAL INFORMATION:
; APPLICANT: N. Alexandrov et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-709P
; CURRENT APPLICATION NUMBER: US/09/513,996A
; CURRENT FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 81028
; SEQ ID NO 67235
; LENGTH: 1114
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 1..1114
; OTHER INFORMATION: any n or xaa - unknown
```

FEATURE:
: OTHER INFORMATION: Location 1..1114 / Ceres Seq. ID 217597
US-09-513-996a-67235

Query Match 99.0%; Score 1039; DB 19; Length 1114;
Best Local Similarity 100.0%; Pred. No. 4,1e-283;
Matches 1039; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 11 TAAATTTTATTTTGTTCAGTAAAGAAATGTCATGTCATTAATGATGCGAGCGC 70
DB 1 TAAATTTTATTTTGTTCAGTAAAGAAATGTCATGTCATTAATGATGCGAGCGC 60
QY 71 TTTTACAGAGAGAGAGCTTCGCTATGAGCGCCATGAGAGCTTCTACATTTCTCCG 130
DB 61 TTTTACAGAGAGAGAGCTTCGCTATGAGCGCCATGAGAGCTTCTACATTTCTCCG 120
QY 131 TTTCTTCTTCTTCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 190
DB 121 TTTCTTCTTCTTCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
QY 191 CCAAGAGAGAGATCCAAATCCCAATTAATGCTGCTGCTGCTGCTGCTGCTGCTG 250
DB 181 CCAAGAGAGAGATCCAAATCCCAATTAATGCTGCTGCTGCTGCTGCTGCTGCTG 240
QY 251 CACTTCTTCTACATCTTATTAAGAGATTTCTCTCTCCGTTGCGAGCTAGAGATTTCC 310
DB 241 CACTTCTTCTACATCTTATTAAGAGATTTCTCTCTCCGTTGCGAGCTAGAGATTTCC 300
QY 311 AGCAATGAGTTGGTGGTGGATTTTGTATGACAGAGCTGCTAGTGGATTTGATTAACG 370
DB 301 AGCAATGAGTTGGTGGTGGATTTTGTATGACAGAGCTGCTAGTGGATTTGATTAACG 360
QY 371 AGCAATGAGTTATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 430
DB 361 AGCAATGAGTTATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
QY 431 TGGCTGGATTTGAGTGGAGACCAATTTTAAAGAGTTCCTCCGGGAGAGTTGC 490
DB 421 TGGCTGGATTTGAGTGGAGACCAATTTTAAAGAGTTCCTCCGGGAGAGTTGC 480
QY 491 GACTGTCAGATATGCTGATAGTAAATTTTGTATGATGTTATGATATTTTCAATTCG 550
DB 481 GACTGTCAGATATGCTGATAGTAAATTTTGTATGATGTTATGATATTTTCAATTCG 540
QY 551 TATCCCTCAAGTCTCCAAATCATTTATGAAGATTCACACTTGTAAAGTATGAGTATGG 610
DB 541 TATCCCTCAAGTCTCCAAATCATTTATGAAGATTCACACTTGTAAAGTATGAGTATGG 600
QY 611 AATTGATGATGATCTGTGAGAGCTTTTCCATGATGATGATGATGATGATGATGATG 670
DB 601 AATTGATGATGATCTGTGAGAGCTTTTCCATGATGATGATGATGATGATGATGATG 660
QY 671 GGATCTTTCAGATTTTACCAACCAAAATTTGTTGAGATTAATAATGAGGCTTTGCTC 730
DB 661 GGATCTTTCAGATTTTACCAACCAAAATTTGTTGAGATTAATAATGAGGCTTTGCTC 720
QY 731 ACTAATGAGAGAGCTTTTTCGAAAGAGCTCTGTAAGCCAAACGAATCAGCTTTGGAA 790
DB 721 ACTAATGAGAGAGCTTTTTCGAAAGAGCTCTGTAAGCCAAACGAATCAGCTTTGGAA 780
QY 791 CTGGAGATTTATCTCTGTCGAAAGAGAGCTTATGAGAGAGAGAGAGAGAGAGAGAG 850
DB 781 CTGGAGATTTATCTCTGTCGAAAGAGAGAGCTTATGAGAGAGAGAGAGAGAGAGAG 840
QY 851 TTCAATGAGATCTTATCAAGGTTCTTAAGAGAGCTTCTGATGCTGCTGAGTCTTAAAG 910
DB 841 TTCAATGAGATCTTATCAAGGTTCTTAAGAGAGCTTCTGATGCTGCTGAGTCTTAAAG 900
QY 911 TGAAGAGAGAGAGCTTAAAGTTAGCCCTATTAACCCAGAGAGTATGATCAATGATATG 970
DB 901 TGAAGAGAGAGAGCTTAAAGTTAGCCCTATTAACCCAGAGAGTATGATCAATGATATG 960
QY 971 ACACCTAATCTAGTCAAGTATGATGATCTTGTGATATTTGATCTAGTCTGCTGCTCCT 1030
```

```

DB 961 ACACCTAATCTAGTCAAGTATGATGATCTTGTGATATTTGATCTAGTCTGCTGCTCCT 1020
QY 1031 TTAACCTCCAGAACTAG 1049
DB 1021 TTAACCTCCAGAACTAG 1039
```

RESULT 3

US-09-513-996a-28855
Sequence 28855, Application US/09513996a

GENERAL INFORMATION:

APPLICANT: N. ALEXANDROV et al.

TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID

TITLE OF INVENTION: ENCODED THEREBY

FILE REFERENCE: 2750-709P

CURRENT APPLICATION NUMBER: US/09/513.996a

CURRENT FILING DATE: 2000-02-25

NUMBER OF SEQ ID NOS: 81028

SEQ ID NO 28855

LENGTH: 1129

TYPE: DNA

ORGANISM: Arabidopsis thaliana

FEATURE:

NAME/KEY: UNSURE

LOCATION: 1..1129

OTHER INFORMATION: any n'or Xaa = unknown

FEATURE:

OTHER INFORMATION: Location 1..1129 / Ceres Seq. ID 1572651

US-09-513-996a-28855

Query Match

Best Local Similarity 95.5%; Score 1002.2; DB 19; Length 1129;

Best Local Similarity 97.8%; Pred. No. 1.1e-272;
Matches 1016; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

```

QY 11 TAAATTTTATTTTGTTCAGTAAAGAAATGTCATGTCATTAATGATGCGAGCGC 70
DB 2 TTAATTTGTTTGTTCATTAAGAAATGTCATGTCATTAATGATGCGAGCGC 61
QY 71 TTTTACAGAGAGAGAGCTTCGCTATGAGCGCCATGAGAGCTTCTACATTTCTCCG 130
DB 62 TTTTACAGAGAGAGAGCTTCGCTATGAGCGCCATGAGAGCTTCTACATTTCTCCG 121
QY 62 TTTTACAGAGAGAGAGCTTCGCTATGAGCGCCATGAGAGCTTCTACATTTCTCCG 121
DB 131 TTTCTTCTTCTTCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 190
QY 131 TTTCTTCTTCTTCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 190
DB 122 TTTCTTCTTCTTCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 181
QY 191 CCAAGAGAGAGATCCAAATCCCAATTAATGCTGCTGCTGCTGCTGCTGCTGCTG 250
DB 182 CCAAGAGAGAGATCCAAATCCCAATTAATGCTGCTGCTGCTGCTGCTGCTGCTG 241
QY 251 CACTTCTTCTACATCTTATTAAGAGATTTCTCTCCGTTGCGAGCTAGAGATTTTCC 310
DB 242 CACTTCTTCTACATCTTATTAAGAGATTTCTCTCCGTTGCGAGCTAGAGATTTTCC 301
QY 311 AGCAATGAGTTGGTGGTGGATTTTGTATGAGAGAGCTCTACTAGTGGATTAACG 370
DB 302 AGCAATGAGTTGGTGGTGGATTTTGTATGAGAGAGCTCTACTAGTGGATTAACG 361
QY 371 AGCAATGAGTTTATTAAGTTCTGATACCAAGAGAGATTAATCTGGAATAGCTTTTGT 430
DB 362 AGCAATGAGTTTATTAAGTTCTGATACCAAGAGAGATTAATCTGGAATAGCTTTTGT 421
QY 431 TGGCTTGATATTTAGTGGAGACAGTTTATGAAGAGTTCCTCCGGGAGAGTTGC 490
DB 422 TGGCTTGATATTTAGTGGAGACAGTTTATGAAGAGTTCCTCCGGGAGAGTTGC 481
QY 491 GACTGTCAGATATGCTGATAGTAAATTTTGTATGATGTTATGATATTTTCAATTCG 550
DB 482 AACTGTCAGATATGCTGATAGTAAATTTTGTATGATGTTATGATATTTTCAATTCG 541
QY 551 TATCCCTCAAGTCTCCAAATCATTTATGAAGATTCACACTTGTAAAGTATGATTTGG 610
```

```

Db      542 TATCCCTCAAGCTCCACATCTTATTGAGATTCCACACTTGTAAAGTGATTTGG 601
Qy      611 AATTGATGGTGCATCTGTGAACCTTTTCCATGACTATGAGTATGATCAAGATTTGA 670
        |||
Db      602 AATTGATGGTGCATCTGTGAACCTTTTCCATGACTATGAGTATGATCAAGATTTGA 661
Qy      671 GGATTTTCAGATTTCAGCCAAACCAAAATTTGGTGAGTAAATAATGGGCTTGGCTTC 730
        |||
Db      662 GGATTTTCAGATTTCAGCCAAACCAAAATTTGGTGAGTAAATAATGGGCTTGGCTTC 721
Qy      721 ACTATGAGACATCTGTTTGAAGAGTCTCTGAAGCCAAACAGATATAGGCTTGGGAA 790
        |||
Db      722 ACTATGAGACATCTGTTTGAAGAGTCTCTGAAGCCAAACAGATATAGGCTTGGGAA 781
Qy      791 CTGGAGATTTCCTCTGTCAAGACAGAGTATACATACGACAGAGATGCTTATGC 850
        |||
Db      782 CTGGAGATTTCCTCTGTCAAGACAGAGTATACATACGACAGAGATGCTTATGC 841
Qy      851 TTGATGCACTCTTTCAGAGGTTCTTAAGACCTTCCTGATGCTGATGCTCATACG 910
        |||
Db      842 TTGATGCACTCTTTCAGAGGTTCTTAAGACCTTCCTGATGCTGATGCTCATACG 901
Qy      911 TGAAGGAGAGCTTAAAGGTTAGCTATACCCCAAGGTTAGATCAATGATATGAT 970
        |||
Db      902 TGAAGGAGAGCTTAAAGGTTAGCTATACCCCAAGGTTAGATCAATGATATGAT 961
Qy      971 ACACCTAACTAGTCAAGTATGATGCAATTCCTGATATGATATGATGCTTGCCT 1030
        |||
Db      962 ACACCTAACTAGTCAAGTATGATGCAATTCCTGATATGATATGATGCTTGCCT 1021
Qy      1031 TTAACCGTCCAGAACTAG 1049
        |||
Db      1022 TTAATCGTCCAGAACTAG 1040

RESULT 4
US-09-708-427-27137
; Sequence 27137, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 27137
; LENGTH: 942
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: 1..942
; OTHER INFORMATION: any n = a, g, c, t, unknown, or other
; NAME/KEY: misc-feature
; LOCATION: 1..942
; OTHER INFORMATION: Ceres Seq. ID 1820252
US-09-708-427-27137

Query Match      79.2%; Score 831; DB 28; Length 942;
Best Local Similarity 96.1%; Pred. No. 3,6e-224;
Matches 852; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

```

```

Db      121 ACCGTACAAGCTACAACTCCCTCATGGCCACAGAGAGATCCAAATCAATCCCAAT 180
Qy      222 AATATCCGTGGCAATTTGGCTGTTCATCAGCTCTTCTACATCTTATTAAGATTTCT 281
        |||
Db      181 AATATCCGTGGCAATTTGGCTGTTCATCAGCTCTTCTACATCTTATTAAGATTTCT 240
Qy      282 CTCTCCGTTGGCCAGCTAGAAATTTCCAGCAATGAGTTGGTGATGATTTGTAT 341
        |||
Db      241 CTCTCCGTTGGCCAGCTAGAAATTTCCAGCAATGAGTTGGTGATGATTTGTAT 300
Qy      342 AGCAAGACTGTCTACGAGTTGATTAAGCAGCAATGACCTTATTAAGTTCTTGATAC 401
        |||
Db      301 AGCAAGACTGTCTACGAGTTGATTAAGCAGCAATGACCTTATTAAGTTCTTGATAC 360
Qy      402 AAGAGATGAATTCGGAATAGCTTTTGTGGCTGGATTAATGAGTGAGACCAAGTTT 461
        |||
Db      361 AAGAGATGAATTCGGAATAGCTTTTGTGGCTGGATTAATGAGTGAGACCAAGTTT 420
Qy      462 AGAAAAGTGTCTCCCGGGAAGTTGGCACTGTCCAGATATGTGTAGATATATAT 521
        |||
Db      421 AGAAAAGTGTCTCCCGGGAAGTTGGCACTGTCCAGATATGTGTAGATATATAT 480
Qy      522 TGTGATGTATGATATTTTTCATTCGTGTATCCCTCAAGTCTCCAACTTTATGAA 581
        |||
Db      481 TGTGATGTATGATATTTTTCATTCGTGTATCCCTCAAGTCTCCAACTTTATGAA 540
Qy      582 GATTCACACTGTGAAGGTAGTATTTGAATGTAGTGTAGTGTGAAGCTTTTCCAT 641
        |||
Db      541 GATTCACACTGTGAAGGTAGTATTTGAATGTAGTGTAGTGTGAAGCTTTTCCAT 600
Qy      642 GACTATGATAGTATATCAAGATGTGAGATCTTTCAGATTAACCAACCAAAATAT 701
        |||
Db      601 GACTATGATAGTATATCAAGATGTGAGATCTTTCAGATTAACCAACCAAAATAT 660
Qy      702 GGTGAGATTAATAATGGGCTTGGCTGTCTACTGACATCTTGTGCAAGAGCTC 761
        |||
Db      661 GGTGAGATTAATAATGGGCTTGGCTGTCTACTGACATCTTGTGCAAGAGCTC 720
Qy      762 CTGAAGCCAAACAGATCAGGCTTGGGAACCTGGAGTTTATCTCTGTCAAGCAGAC 821
        |||
Db      721 CTGAAGCCAAACAGATCAGGCTTGGGAACCTGGAGTTTATCTCTGTCAAGCAGAC 780
Qy      822 TTACAAATACGACCAACGAGATCTTATGCTTATGATGATCTTTTAAAGCTTT 881
        |||
Db      781 TTACAAATACGACCAACGAGATCTTATGCTTATGATGATCTTTTAAAGCTTT 840
Qy      882 CTTCCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 928
        |||
Db      841 AAAAACCATCTTCTCACACTCAACGACCTTGAAGCAAAAATCTCACAA 887

RESULT 5
US-09-896-186b-1
; Sequence 1, Application US/09896186B
; GENERAL INFORMATION:
; APPLICANT: Joshua Z. Levin
; APPLICANT: Ken Phillips
; APPLICANT: Greg Budziszewski
; APPLICANT: Fred Meins
; APPLICANT: Zhenya Glazov
; TITLE OF INVENTION: Methods of Controlling Gene Expression
; FILE REFERENCE: PB/5-31481A
; CURRENT APPLICATION NUMBER: US/09/896,186B
; CURRENT FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 1
; LENGTH: 942
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-896-186b-1

```


Query Match 79.2%; Score 831; DB 33; Length 942;
Best Local Similarity 96.1%; Pred. No. 3,66-224;
Matches 852; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 42 ATGTCATCGTCAAAATTTGGATCGACGACGCTTTTACAGAGAGAGCTTTCGCTATGAC 101
DB 1 ATGTCATCGTCAAAATTTGGATCGACGACGCTTTTACAGAGAGAGCTTTCGCTATGAC 60
QY 102 GCCATCGACGCTTCTCAAAATTTTCGCGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 161
DB 61 GCCATCGACGCTTCTCAAAATTTTCGCGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
QY 162 ACCGTACAGCTACAACTCCGCTCCATGAGCAGAGAGAGTCAATCAATCCCAAT 221
DB 121 ACCGTACAGCTACAACTCCGCTCCATGAGCAGAGAGAGTCAATCAATCCCAAT 180
QY 222 AATATCGGTGCCAATTTGCTCTGTCATCACTTCTCTACATCTTAAACGATTTCT 281
DB 181 AATATCGGTGCCAATTTGCTCTGTCATCACTTCTCTACATCTTAAACGATTTCT 240
QY 282 CTCTCCGCTTGCAGCTAGGAATTTTCAGCAATGAGTTGGTGGATGATTTGTAT 341
DB 241 CTCTCCGCTTGCAGCTAGGAATTTTCAGCAATGAGTTGGTGGATGATTTGTAT 300
QY 342 AGCAAGACTCTACTGAGTTGATTAAGCAGCAATGAGCTTAAAGTCTTGATACC 401
DB 301 AGCAAGACTCTACTGAGTTGATTAAGCAGCAATGAGCTTAAAGTCTTGATACC 360
QY 402 AAGAGATGAATCTGGAATAGCTTTTGTGGCTTGATATGAGTGAGACCAAGTTT 461
DB 361 AAGAGATGAATCTGGAATAGCTTTTGTGGCTTGATATGAGTGAGACCAAGTTT 420
QY 462 AGAAAGGTCTTCCCGGGGAGGTTGCGACTGTCAGANATGTAGATAGATATAT 521
DB 421 AGAAAGGTCTTCCCGGGGAGGTTGCGACTGTCAGANATGTAGATAGATATAT 480
QY 522 TGTGATGTATGATATTTTTCATCTGATCCCTCAAGTCCCAACATCTTATGAA 581
DB 481 TGTGATGTATGATATTTTTCATCTGATCCCTCAAGTCCCAACATCTTATGAA 540
QY 582 GATTCAACACTTGTAAAGATGATTTGGAATGATGATGATGATGATGATGATGAT 641
DB 541 GATTCAACACTTGTAAAGATGATTTGGAATGATGATGATGATGATGATGATGAT 600
QY 642 GACTATGATGATGATCAAGATGATGATGATGATGATGATGATGATGATGATGAT 701
DB 601 GACTATGATGATGATCAAGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 702 GGTGAGATTAATAAATGGGCTTGCCTCACTAATGAGACACTTGTGCAAGAGCTC 761
DB 661 GGTGAGATTAATAAATGGGCTTGCCTCACTAATGAGACACTTGTGCAAGAGCTC 720
QY 762 CTGAGCAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 821
DB 721 CTGAGCAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 822 TTACATATGAGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 881
DB 781 TTACATATGAGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 882 CTTCCTGATGCTGCTACTGCTCATATAGCTGAAGAGAGAGCTTAA 928
DB 841 AAAAAGCATCTTCTCAACATCAAGACCTTGAAGCAAAATCTCACA 887

RESULT 6
US-09-620-393B-6112
; Sequence 6112, Application US/09620393B
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1068P

CURRENT APPLICATION NUMBER: US/09/620,393B
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 9948
SOFTWARE: PatentIn version 3.0
SEQ ID NO 6112
LENGTH: 507
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc-feature
LOCATION: 1..507
OTHER INFORMATION: any n = a, g, c, t, unknown, or other
FEATURE:
NAME/KEY: misc-feature
LOCATION: 1..507
OTHER INFORMATION: Ceres Seq. ID 1392052
US-09-620-393B-6112

Query Match 45.0%; Score 471.6; DB 24; Length 507;
Best Local Similarity 97.0%; Pred. No. 2,1e-122;
Matches 491; Conservative 0; Mismatches 14; Indels 1; Gaps 1;

QY 11 TAAATTTTATTTTGTGTTTCAGTAAAGAAATGTCATCGTCAAAATGATGACGACGC 70
DB 2 TTAATTTGTTTGTGTTTCATTAAGAAATGTCATCGTCAAAATGATGACGACGC 61
QY 71 TTTTACAGAGAAAGCTTCGCTATGACGACGATGAGCTTCTTCAATTTCTCCG 130
DB 62 TTTTACAGAGAAAGCTTCGCTATGACGACGATGAGCTTCTTCAATTTCTCCCT 121
QY 131 TTTCTTCT 190
DB 122 TTTCTTCT 181
QY 191 CCACAGAGAGATCCAAATCAATCCCAATTAATATGCTGCGCAATTTGCTGTCAT 250
DB 182 CCACAGAGAGATCCAAATCAATCCCAATTAATATGCTGCGCAATTTGCTGTCAT 241
QY 251 CACTTCTCTCACTTATTAAGATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 310
DB 242 CACTTCTCTCACTTATTAAGATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 301
QY 311 AGCATGAGGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 370
DB 302 AGCATGAGGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 361
QY 371 AGCAATGAGCTTAAAGTCTTGAATACCAAGAGATGAATCTGGAATAGCTTTGT 430
DB 362 AGCAATGAGCTTAAAGTCTTGAATACCAAGAGATGAATCTGGAATAGCTTTGT 421
QY 431 TGGCTTGAATTTAGTGAAGACCAAGTTTAAAGAGTCTCTCCGAGGAA-GGTG 489
DB 422 TGGCTTGAATTTAGTGAAGACCAAGTTTAAAGAGTCTCTCCGAGGAAAGGTG 481
QY 490 CGACTGTCCAGATATGTTAGTAGT 515
DB 482 CAACTGTCCAGATATGTTAGTAGTAAT 507

RESULT 7
US-09-534-859-579/c
; Sequence 579, Application US/09534859
; GENERAL INFORMATION:
; APPLICANT: Bush, David F.
; APPLICANT: Last, Robert L.
; APPLICANT: Levin, Irena M.
; APPLICANT: Norris, Susan R.
; APPLICANT: Parnell, Laurence D.
; APPLICANT: Rounsley, Steven D.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: PLANT POLYMORPHIC MARKERS AND USES THEREOF
; FILE REFERENCE: 38-10(15493)B
; CURRENT APPLICATION NUMBER: US/09/534,859

;; CURRENT FILING DATE: 2000-03-29
;; NUMBER OF SEQ ID NOS: 1127
;; SEQ ID NO 579
;; LENGTH: 118718
;; TYPE: DNA
;; ORGANISM: Arabidopsis thaliana
US-09-534-859-579

Query Match 28.8%; Score 302.2; DB 20; Length 118718;
Best Local Similarity 97.5%; Pred. No. 3.5e-73;
Matches 307; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 CCAAGCATTAATTTTATTTTGTTCAGTAAAGAAATGTCATCGTCAATTGGAT 61
DB 96598 CCAAGCATTAATTTTATTTTGTTCAGTAAAGAAATGTCATCGTCAATTGGAT 96539
QY 62 CGACGAGCTTTTACAGAGAGAGAGCTTCGCTATCGAGCGCATGGAAGCTTCTACAA 121
DB 96538 CGACGAGCTTTTACAGAGAGAGAGCTTCGCTATCGAGCGCATGGAAGCTTCTACAA 96479
QY 122 TTTTCGCCGTTCTTCTTCTTCTTCTTCTGCTGCTCGACCGTACAGCTACAACTTC 181
DB 96478 TTTTCGCCGTTCTTCTTCTTCTTCTTCTGCTGCTCGACCGTACAGCTACAACTTC 96419
QY 182 CGTCATGGCCACGAGAGAGATCCAAATCCCAATATATCCGTCGCAATTGGC 241
DB 96418 CGTCATGGCCACGAGAGAGATCCAAATCCCAATATATCCGTCGCAATTGGC 96359
QY 242 TCGTTCATCACTTCTTCTACATCTTATTAAGATTTCTCTCTCCGTTGCGAGCTAG 301
DB 96358 TCGTTCATCACTTCTTCTACATCTTATTAAGATTTCTCTCTCCGTTGCGAGCTAT 96299
QY 302 GAATTTCCAGCAAT 316
DB 96298 GAACCTTCTTCAAT 96284

RESULT 8
US-09-803-736-579/c
; Sequence 579, Application US/09803736
; GENERAL INFORMATION:
; APPLICANT: Bush, David F.
; APPLICANT: Levin, Irena M.
; APPLICANT: Norris, Susan R.
; APPLICANT: Rounsley, Steven D.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Plant Polymorphic Markers and Uses Thereof
; FILE REFERENCE: 38-10(15493)D
; CURRENT APPLICATION NUMBER: US/09/803,736
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 09/534,859
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 1582
; SEQ ID NO 579
; LENGTH: 118718
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-803-736-579

Query Match 28.8%; Score 302.2; DB 31; Length 118718;
Best Local Similarity 97.5%; Pred. No. 3.5e-73;
Matches 307; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 CCAAGCATTAATTTTATTTTGTTCAGTAAAGAAATGTCATCGTCAATTGGAT 61
DB 96598 CCAAGCATTAATTTTATTTTGTTCAGTAAAGAAATGTCATCGTCAATTGGAT 96539
QY 62 CGACGAGCTTTTACAGAGAGAGCTTCGCTATCGAGCGCATGGAAGCTTCTACAA 121
DB 96538 CGACGAGCTTTTACAGAGAGAGCTTCGCTATCGAGCGCATGGAAGCTTCTACAA 96479

QY 122 TTTTCGCCGTTCTTCTTCTTCTTCTGCTGCTCGACCGTACAGCTACAACTTC 181
DB 96478 TTTTCGCCGTTCTTCTTCTTCTTCTGCTGCTCGACCGTACAGCTACAACTTC 96419
QY 182 CGTCATGGCCACGAGAGAGATCCAAATCCCAATATATCCGTCGCAATTGGC 241
DB 96418 CGTCATGGCCACGAGAGAGATCCAAATCCCAATATATCCGTCGCAATTGGC 96359
QY 242 TCGTTCATCACTTCTTCTACATCTTATTAAGATTTCTCTCTCCGTTGCGAGCTAG 301
DB 96358 TCGTTCATCACTTCTTCTACATCTTATTAAGATTTCTCTCTCCGTTGCGAGCTAT 96299
QY 302 GAATTTCCAGCAAT 316
DB 96298 GAACCTTCTTCAAT 96284

RESULT 9
US-09-906-226-15
; Sequence 15, Application US/09906226
; GENERAL INFORMATION:
; APPLICANT: Butler, Karla
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Rafalski, Antoni
; APPLICANT: Sakai, Hajime
; TITLE OF INVENTION: Plant RNaseD-Like Genes
; FILE REFERENCE: B01467 US NA
; CURRENT APPLICATION NUMBER: US/09/906,226
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 60/218993
; PRIOR FILING DATE: July 17, 2000
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 1170
; TYPE: DNA
; ORGANISM: Glycine max
US-09-906-226-15

Query Match 26.3%; Score 276.2; DB 34; Length 1170;
Best Local Similarity 66.6%; Pred. No. 8.2e-67;
Matches 395; Conservative 0; Mismatches 198; Indels 0; Gaps 0;

QY 300 AGCAATTTTCCAGCAATGAGGTTGGTGTAGATTTTGTATGACCAAGCTGCTACTGAG 359
DB 229 ATGACATTTCTCTGTATGAAAGTTAGTGTCAAAATTTCTTATAGCAGCACTTTGTATGCT 288
QY 360 GTTGATTAACGAGCAATGAGCTTATTAAGTCTTGATACCAAGAGATGATGTGA 419
DB 289 GTAGAGAAAGCTGCAAAAGCTTTCACAAATTTCCAGAAAAGACCGACATGATG 348
QY 420 ATGACTTTTGGCTTGGATATTTGAGTGAGACCAAGTTTGTAGAAAAGGTGTCGCG 479
DB 349 CAACCTGCAATTTGATTTGACATTTGAGTGGAAAACCACTTCAGAAAAGGTGTCGCG 408
QY 480 GGAAGTTGCGACTGTCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 539
DB 409 GGAAGTTGCGACTGTCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 468
QY 540 TTTCAATTTGATTTCTCTCAAAAGTCTCCAACTCTTATGAAAGTTCAACCTGTAAG 599
DB 469 ATTCATTTGGAATCCCTCAAAATTTACAGCTTTGCTTGAAGATCCACAGCTTGAAG 528
QY 600 GTAGATTTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 659
DB 529 GTTGGAGCTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 588
QY 660 AAAGATTTGAGATTTTCAATTTAGCAACCAAAATTTGTGAGATGATGATGATGATGAT 719
DB 589 AAAGATTTGAGATTTTCTTTCTTCACTTATTAAGAACTTGTGAGATGATGATGATGAT 648
QY 720 GGCCTTTCCTCACTAAGTGAACACTTTGTTGCAAAAGAGCTCTGTAAGCAACAGATTC 779

Db 649 GGTCTTGACATCTTGACTGAAACAACTTCTATCAAAACAGCTTAAGAACCCCAACAAATA 708
QY 780 AGGCTTGGAGACTGGAGTTTATCCCTCTCAAGAGACAGATTACATAGCGACAGC 839
Db 709 AGACTGGAAATGGAGGCTCCCTGTTTCAAGAGACACTAGATGTGTCACACA 768
QY 840 GATGCTTATGCTTCATGCACTTTTACAAAGTTCTTTAGGACCTTCTGTATGC 892
Db 769 GATGCTTATGCTTCCTTGTGTCTTATCAAGCGCATTAAGATCTCCCGGACGC 821

RESULT 10
US-60-218-993-15
; Sequence 15, Application US/60218993
; GENERAL INFORMATION:
; APPLICANT: Butler, Karla
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Rafalski, Antoni
; TITLE OF INVENTION: Plant RNased-Like Genes
; FILE REFERENCE: B01467 US PRV
; CURRENT APPLICATION NUMBER: US/60/218,993
; CURRENT FILING DATE: 2000-07-17
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 1170
; TYPE: DNA
; ORGANISM: Glycine max
US-60-218-993-15

Query Match 26.3%; Score 276.2; DB 65; Length 1170;
Best Local Similarity 66.6%; Pred. No. 8.2e-67;
Matches 395; Conservative 0; Mismatches 198; Indels 0; Gaps 0;

QY 300 AGCAATTTTCCACCAATGAGCTTGTGTAGAGATTTTGTATGCAAGACTGCTACTGAG 359
Db 229 ATACATTTGCTCTTAATGAAGTTAGTCAAAATTTCTTATAGCAGACTTTTGAGTCT 288
QY 360 GTTGATTAAGGAGCAATGACGCTTATTAAGTTCTTGATACCAAGAGATGATCGGA 419
Db 289 GTTGAGAAAGCTGCACAAAGCTCTTCAAAATTTCTCCAGAAAAGAGACCGACATGATG 348
QY 420 ATAGCTTTTGTGCTGATTTGATTTGAGAGACCAAGTTTGTAGAAAAGTGTCTCCG 479
Db 349 CAACAGCAATTTGATTTGACATTTGAGTGAAGAACCCCTTCAGAAAAGTGTCTCCG 408
QY 480 GGAAGAGTTCGAGCTGTCAGATATGTATGATATGATATTAATTTGTGTATGATGAT 539
Db 409 GGAAGAGTTCGAGCTGATGATGATATGATGATGATGATGATGATGATGATGATGAT 468
QY 540 TTTCATTTGCTGATTCCTCAAAAGTCTCCACATCTTATTAAGATTTCAACATTTGAAG 599
Db 469 ATTCATTTGGAATCCCTCAAAATTTACAGCTTTTGTGAAGATCCCAAGCTTTGAAG 528
QY 600 GTAGTATTTGGAATTTGATGCTGCTGAGAGCTTTTCCATGACTTTGAGAGTATGATC 659
Db 529 GTTGAGAGCTGGATTTGATGCTGATGCTGATGAGTTTGTAGAGATTTTAAATATCTG 588
QY 660 AAAGATTTGAGAGATCTTTCAGATTTAGCCAAACCAAAATTTGGTGGAGATTAAGAA 719
Db 589 AAAGATTTGAGAGATCTTTCATGATTTATCAAAAGCTTTGGTGGAGATTAAGAA 648
QY 720 GGCTTGGCTCATCTACTGACACTGTTTGAAGAGAGTCTCGAAGCCCAACAGATC 779
Db 649 GGCTTGGCTCATCTTGAAGTGAAGAACTTTCATCAAAAGCTTTAAGAAAGCCCAACAAATA 708
QY 780 AGGCTTGGAGACTGGAGTTTATCCCTCTCAAGAGACAGATTACATAGCGACAGC 839
Db 709 AGACTGGAAATGGAGGCTCCCTGTTTCAAGAGACACTAGATGTGTCACACA 768
QY 840 GATGCTTATGCTTCATGCACTTTTACAAAGTTCTTTAGGACCTTCTGTATGC 892
Db 769 GATGCTTATGCTTCCTTGTGTCTTATCAAGCGCATTAAGATCTCCCGGACGC 821

RESULT 11
US-09-605-9441
; Sequence 9441, Application US/09605698
; GENERAL INFORMATION:
; APPLICANT: Havukkala, Ilkka
; APPLICANT: Sheen, Michael A.
; TITLE OF INVENTION: Polynucleotides, Materials Incorporated
; FILE REFERENCE: 11000.104101
; CURRENT APPLICATION NUMBER: US/09/605,698
; CURRENT FILING DATE: 2000-06-21
; NUMBER OF SEQ ID NOS: 25120
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9441
; LENGTH: 672
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-605-9441

Query Match 17.3%; Score 181.2; DB 23; Length 672;
Best Local Similarity 60.9%; Pred. No. 5.4e-40;
Matches 313; Conservative 0; Mismatches 198; Indels 3; Gaps 1;

QY 382 TTATTAAGTTCTTGATACCAAGAGATGATGATGATGATGATGATGATGATGATGATG 441
Db 1 TTTTGAAGAAATCAATGCAATGAAAGTGTATGATCAAGTCAATATATGCTTTGACA 60
QY 442 TTGAGTGAGACCAAGTTTGAAGAAAGTGTCTCCGAGGAGGTTGCGACTGTCAGA 501
Db 61 TTGAGTATGAGCCCAATTTCAAAAGTGTCTTCAACTCGAGAGACGCAATATATGAGA 120
QY 502 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 561
Db 121 TTGTGATGACAAATGACACCTGATGATGATGATGATGATGATGATGATGATGATGAT 180
QY 562 GTCTCCAAACCTTATGAAAGTCAACACTTGTAAAGGATGATGATGATGATGATGATG 621
Db 181 GTTGAAGCTTATTTAGAGATCAACTGATATTAAGAGTGTGAGAGGCAATGATGATG 240
QY 622 ACTCTGTAAGCTTTTCCATGATGATGATGATGATGATGATGATGATGATGATGATG 681
Db 241 ATGTAGGAAGGCTCTTAAGATTTACATGATGATGATGATGATGATGATGATGATGAT 300
QY 682 ATTTAGCCAAACCAAAATG--GTGAGATTAAGAAATGAGGCTTGTGCTACTAATG 738
Db 301 AACTTGCTTAATCAAAAGCTTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
QY 739 AGACACTTGTGCAAGAGCTCTGCAAGCAACAAATCAAGCTTGGAACTGGAGAT 798
Db 361 AGTGTCTTCAATATGACAGCTGAACAAATTCAGAAATTAAGGATGAGGAAATGGAGG 420
QY 799 TTTATCTCTGTCAAGCAGATTTACATATGAGCAGACAGATGATGATGATGATGATG 858
Db 421 CTCTTTTATCAAAAGCAACACTTTATGCTGACATGACATGACATGATGATGATGATG 480
QY 859 ATCTTACAAAGCTTTTAAGGACCTTCTGTATGC 892
Db 481 ATCTATATCAAGCTTTAAGGAGCTGCGCAGATGC 514

RESULT 12
US-09-849-529A-2068
; Sequence 2068, Application US/09849529A
; GENERAL INFORMATION:
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-21(51893)B
; CURRENT APPLICATION NUMBER: US/09/849,529A
; CURRENT FILING DATE: 2001-05-07

; PRIOR APPLICATION NUMBER: US 60/196,868
 ; PRIOR FILING DATE: 2000-05-09
 ; NUMBER OF SEQ ID NOS: 24076
 ; SEQ ID NO 2068
 ; LENGTH: 584
 ; TYPE: DNA
 ; ORGANISM: Gossypium hirsutum
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(584)
 ; OTHER INFORMATION: unsure at all n locations
 ; OTHER INFORMATION: Clone ID: LIB3543-036-Q1-K1-P5
 US-09-849-529A-2068

Query Match 16.7%; Score 174.8; DB 32; Length 584;
 Best Local Similarity 64.4%; Pred. No. 3.3e-38;
 Matches 282; Conservative 0; Mismatches 147; Indels 9; Gaps 1;

QY 208 ATCAATCCCAATATATCCGTCGCAATGCGTTCATCCACTTCTTACATCTT 267
 DB 147 ATCAAGATACCTCAAAACACAGCCGCAATGCCCCATCTATCTCACCAGCCGCTT 206
 QY 268 ATAAAGATTTCCCTCTCCGCTGCGGAGCTA-----GGAATTTCCAGCAATGA 318
 DB 207 CTTTCCGCTCTCTCGCTGCTGCTCAATCTAATACAAAGCTTGAATCCGCGTTGA 266
 QY 319 GATTGGTGTAGATTTTGTATACCAAGACTGCTACTGAGGTTGATAGCGCAATGC 378
 DB 267 GATTGGAGGTCATATTTTGTATACCTTTAGCGAAGATGAGAGAAATGCTGCATGCG 326
 QY 379 AGCTATTAAGTCTTGTATACCAAGAGATGAATCGGAATGCTTTTGTGCTTGG 438
 DB 327 AGCTATTAAGTCTTGTATACCAAGAGATGAATGCTTTGCTTTAGGAGTTTG 386
 QY 439 ATATTGAGTGGAGCAAGATTTTGAAGAGGTTCTCCGCGGAGGTTGCGACGTCC 498
 DB 387 ATATTGAGTGGAGCAAGATTTTGAAGAGGTTCTCCGCGGAGGTTGCGGATGCG 446
 QY 499 AGATATGCTGATAGTATATATGTTGATGATGATGATGATGATGATGATGATG 558
 DB 447 AGATATGCTGATAGTATATATGTTGATGATGATGATGATGATGATGATGATG 506
 QY 559 AAAGCTCCCAATCTTATGTAAGATTCACACTTGAAGTAGTATGGAATTTGATG 618
 DB 507 AAAGCTCCCAATCTTATGTAAGATTCACACTTGAAGTAGTATGGAATTTGATG 566
 QY 619 GTGACTGTGTAAGCTTT 636
 DB 567 GCGATGCTATCAAGGTGT 584

RESULT 13
 US-60-196-868-2068
 ; Sequence 2068, Application US/60196868
 ; GENERAL INFORMATION:
 ; APPLICANT: Fincher, Karen L.
 ; APPLICANT: McCarter, David W.
 ; APPLICANT: Ziegler, Todd E.
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
 ; FILE REFERENCE: PLANTS
 ; CURRENT APPLICATION NUMBER: 38-21(51893)A
 ; CURRENT FILING DATE: US/60/196,868
 ; NUMBER OF SEQ ID NOS: 24054
 ; SEQ ID NO 2068
 ; LENGTH: 584
 ; TYPE: DNA
 ; ORGANISM: Gossypium hirsutum
 ; OTHER INFORMATION: Clone ID: LIB3543-036-Q1-K1-P5
 US-60-196-868-2068

Query Match 16.7%; Score 174.8; DB 63; Length 584;
 Best Local Similarity 64.4%; Pred. No. 3.3e-38;

Matches 282; Conservative 0; Mismatches 147; Indels 9; Gaps 1;

QY 208 ATCAATCCCAATATATCCGTCGCAATGCGTTCATCCACTTCTTACATCTT 267
 DB 147 ATCAAGATACCTCAAAACACAGCCGCAATGCCCCATCTATCTCACCAGCCGCTT 206
 QY 268 ATAAAGATTTCCCTCTCCGCTGCGGAGCTA-----GGAATTTCCAGCAATGA 318
 DB 207 CTTTCCGCTCTCTCGCTGCTGCTCAATCTAATACAAAGCTTGAATCCGCGTTGA 266
 QY 319 GATTGGTGTAGATTTTGTATACCAAGACTGCTACTGAGGTTGATAGCGCAATGC 378
 DB 267 GATTGGAGGTCATATTTTGTATACCTTTAGCGAAGATGAGAGAAATGCTGCATGCG 326
 QY 379 AGCTATTAAGTCTTGTATACCAAGAGATGAATCGGAATGCTTTTGTGCTTGG 438
 DB 327 AGCTATTAAGTCTTGTATACCAAGAGATGAATCGGAATGCTTTTGTGCTTGG 386
 QY 439 ATATTGAGTGGAGCAAGATTTTGAAGAGGTTCTCCGCGGAGGTTGCGACGTCC 498
 DB 387 ATATTGAGTGGAGCAAGATTTTGAAGAGGTTCTCCGCGGAGGTTGCGGATGCG 446
 QY 499 AGATATGCTGATAGTATATATGTTGATGATGATGATGATGATGATGATGATG 558
 DB 447 AGATATGCTGATAGTATATATGTTGATGATGATGATGATGATGATGATGATG 506
 QY 559 AAAGCTCCCAATCTTATGTAAGATTCACACTTGAAGTAGTATGGAATTTGATG 618
 DB 507 AAAGCTCCCAATCTTATGTAAGATTCACACTTGAAGTAGTATGGAATTTGATG 566
 QY 619 GTGACTGTGTAAGCTTT 636
 DB 567 GCGATGCTATCAAGGTGT 584

RESULT 14
 US-09-708-427-65909
 ; Sequence 65909, Application US/09708427
 ; GENERAL INFORMATION:
 ; APPLICANT: N. ALEXANDROV et al.
 ; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
 ; FILE REFERENCE: THEREBY
 ; CURRENT APPLICATION NUMBER: US/09/708,427
 ; CURRENT FILING DATE: 2000-11-09
 ; NUMBER OF SEQ ID NOS: 85364
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 65909
 ; LENGTH: 1434
 ; TYPE: DNA
 ; ORGANISM: Zea mays subsp. mays
 ; FEATURE:
 ; NAME/KEY: misc.feature
 ; LOCATION: 1..1434
 ; OTHER INFORMATION: any n - a, g, c, t, unknown, or other
 ; NAME/KEY: misc.feature
 ; LOCATION: 1..1434
 ; OTHER INFORMATION: Ceres Seq. ID 1929706
 US-09-708-427-65909

Query Match 16.6%; Score 174; DB 28; Length 1434;
 Best Local Similarity 57.4%; Pred. No. 8.6e-38;
 Matches 334; Conservative 0; Mismatches 245; Indels 3; Gaps 1;

QY 309 CCAGCAATGAGGTTGGTGTGAGATTTTGTATACCAAGACTGCTAGAGTTGATAG 368
 DB 376 CAACAATAATCATTTAGCGCGTAGATGTTTCTCAGAGACACTTACAGAGCGGAGAAA 435
 QY 369 CGAGCAATGACACTTATTAAGTCTTGTATACCAAGAGATGAATCTGGAATACCTTT 428
 DB 436 GCTGCAATGACACTTATTAAGTCTTGTATACCAAGAGATGAATCTGGAATACCTTT 492
 QY 429 GTTGCTTGATATTTGAGTGTGAGACCAAGTTTACAAAAGGTGTTCTCCCGGGAAGTT 488

```

Db      493  CTTGGATTTGATCTCGAGTGGAGGCCCTTTCCAGAAAGGAGAACCCCTTGTAAAGSTT 552
Qy      489  GCGAGCTGCCAGATATGTGATAGTAATATTTGTATGATATGATATTTTCATCT 548
Db      553  GCTGTAAATGCAACTATGATGATGAGAAACCTGTGCTATGCTGCAATATGCTACACT 612
Qy      549  GGTATCCCTCAAAAGCTTCCAAACATCTTATGAAAGATTCACACTTGTAAAGGTAGTAT 608
Db      613  GGGGTACCGCCTATACAGAAACCTTTTGGAGATAGTTGTCATTTAAAGTTGGAATA 672
Qy      609  GGAATTTGATGGTACTCTGTGAAGCTTTCCATGACTATGAGATTAGTATCAAGATGTT 668
Db      673  TGCATAGACAAATGATGAGAAATGTTTAAATGATGATGATGCTGTGCTACACACTA 732
Qy      669  GAGATCTTTCAGATTTAGCCAAACCAAAAATTTGTTGAGATTAATAAATGGGCTTCC 728
Db      733  ATGGATTTGTGACCTTGGCAAAATGTCAAAGTTAGCTTCTCCCAAAAGATGAGCCTTCT 792
Qy      729  TCACTTAAGTACACTTGTGTTGGCAAGAGCTCTGGAAGCCAAACAGATAGGCTTGG 788
Db      793  TCTTTAACCAGAAATGATACATGCAAAAGATTTGCCAAAGCCGAGCAATAGATGGG 852
Qy      789  AACTGGAGTTTATCTCTGTCAAGCAGGATTAACATAGCAGCAACGATGCTTAT 848
Db      853  AACTGGAGGCTGATGCTCTCTCCAAACAACTTCAATATGCTGCTACAGATGCTTAT 912
Qy      849  GCTTCATGAGCATCTTTACAAAGGTTCTTAAGACCTTCTGAT 890
Db      913  ATCTCATGTACTTGTATAGAGCCCTGCAACTCTTCCAGAT 954

```

```

RESULT 15
US-09-708-427-83271
; Sequence 83271, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 83271
; LENGTH: 1434
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..1434
; OTHER INFORMATION: any n = a, g, c, t, unknown, or other
; NAME/KEY: misc.feature
; LOCATION: 1..1434
; OTHER INFORMATION: Ceres Seq. ID 1962403
US-09-708-427-83271

```

```

Query Match      16.6%; Score 174; DB 28; Length 1434;
Best Local Similarity 57.4%; Pred. No. 8.6e-38;
Matches 334; Conservative 0; Mismatches 245; Indels 3; Gaps 1;

```

```

Qy      309  CCAGCAATGAGGTTGGTGGAGATTTGTATAGCAAGACTGCTAGCAGGTGATTAAG 368
Db      376  CACAAATATCATTCACAGCGGTAAGATGTTTACTGACAGACCTACAGAGCGGAGAAA 435
Qy      369  CGAGCAATGAGCTTATTAAGTCTTGATACCAAGAGAGATGATCTGGAATAGCTTTT 428
Db      436  GCTGCAACTGACATCTTACCAAAATGAGAGATGAAA---ACCCAGGCCAGTTTCT 492
Qy      429  GTTGGCTTGGATATGAGTGAGACCAAGATTTAGAAAAGGTGTTCTCCGGGGAAGGTT 488
Db      493  CTTGGATTTGATCTCGAGTGGAGGCCCTTTCCAGAAAGGAGAACCACTTGTAAAGSTT 552

```

```

Qy      489  GCGAGCTGCCAGATATGTGATAGTAATATTTGTATGATATGATATTTTCATCT 548
Db      553  GCTGTAAATGCAACTATGATGATGAGAAACCTGTGCTATGCTGCAATATGCTACACT 612
Qy      549  GGTATCCCTCAAAAGCTTCCAAACATCTTATGAAAGATTCACACTTGTAAAGGTAGTAT 608
Db      613  GGGGTACCGCCTATACAGAAACCTTTTGGAGATAGTTGTCATTTAAAGTTGGAATA 672
Qy      609  GGAATTTGATGGTACTCTGTGAAGCTTTCCATGACTATGAGATTAGTATCAAGATGTT 668
Db      673  TGCATAGACAAATGATGAGAAATGTTTAAATGATGATGATGCTGTGCTACACACTA 732
Qy      669  GAGATCTTTCAGATTTAGCCAAACCAAAAATTTGTTGAGATTAATAAATGGGCTTCC 728
Db      733  ATGGATTTGTGACCTTGGCAAAATGTCAAAGTTAGCTTCTCCCAAAAGATGAGCCTTCT 792
Qy      729  TCACTTAAGTACACTTGTGTTGGCAAGAGCTCTGGAAGCCAAACAGATAGGCTTGG 788
Db      793  TCTTTAACCAGAAATGATACATGCAAAAGATTTGCCAAAGCCGAGCAATAGATGGG 852
Qy      789  AACTGGAGTTTATCTCTGTCAAGCAGGATTAACATAGCAGCAACGATGCTTAT 848
Db      853  AACTGGAGGCTGATGCTCTCTCCAAACAACTTCAATATGCTGCTACAGATGCTTAT 912
Qy      849  GCTTCATGAGCATCTTTACAAAGGTTCTTAAGACCTTCTGAT 890
Db      913  ATCTCATGTACTTGTATAGAGCCCTGCAACTCTTCCAGAT 954

```

```

Search completed: June 14, 2003, 09:12:07
Job time : 2843 secs

```


QY	1	ACCAAGCATTAATTTTTTTTTTTTGGTTCGTAAAGAAATGCATCGCAAAATGGA	60
Db	1	ACCAAAAGCATTAATTTTTTTTTTTTGGTTCGTAAAGAAATGCATCGCAAAATGGA	60
QY	61	TCGACGAGCCTTTTACAGAGAGAGACTTTCGTATCGACGCCATCGAAGCTTCCTACA	120
Db	61	TCGACGAGCCTTTTACAGAGAGAGACTTTCGTATCGACGCCATCGAAGCTTCCTACA	120
QY	121	ATTTCCTCCGCTTCTTCTTCTTCTCTCTCGTGCCTCGACACGTAAACCTCAACT	180
Db	121	ATTTCCTCCGCTTCTTCTTCTTCTCTCTCTCGTGCCTCGACACGTAAACCTCAACT	180
QY	181	CGTCATATGGCCACAGAGAGATCCCAATCAAAATCCCAATATATATCGTTCGCAATGC	240
Db	181	CGTCATATGGCCACAGAGAGATCCCAATCAAAATCCCAATATATATCGTTCGCAATGC	240
QY	241	CTCGTTCATCACTTCTTCTTACACTTTAAACGATTTCTCTCCCGTTGCCAGACTA	300

```

Db      241 CTCGTCATCATCTCTCTTACATCTATTAAGATTTCCCTCCCGGTCGACGTA 300
Oy      301 GGAATTTTCCAGCAATGAGTGTGTGAGATTTTGTATAGCAAGACTGCTACTAGG 360
Db      301 GGAATTTTCCAGCAATGAGTGTGTGAGATTTTGTATAGCAAGACTGCTACTAGG 360
Oy      361 TTGATAGCGAGCAATGCAAGCTTATTAAGTTCTTGATACCAAGAGATGATCTGAA 420
Db      361 TTGATAGCGAGCAATGCAAGCTTATTAAGTTCTTGATACCAAGAGATGATCTGAA 420
Oy      421 TAGCTTTTGTGGCTGGATATGAGTGAGAGCAAGTTTGAAGAGGTCTCCCGG 480
Db      421 TAGCTTTTGTGGCTGGATATGAGTGAGAGCAAGTTTGAAGAGGTCTCCCGG 480
Oy      481 GGAAGGTTGCGACTGCTCCAGATATGTAGATAGTATATTTGATGTATGCAATATT 540
Db      481 GGAAGGTTGCGACTGCTCCAGATATGTAGATAGTATATTTGATGTATGCAATATT 540
Oy      541 TTGATTTGCTATCCCTCAAGTCTCCAGCATCTTATTTGAGATTCACACTTGTAAAG 600
Db      541 TTGATTTGCTATCCCTCAAGTCTCCAGCATCTTATTTGAGATTCACACTTGTAAAG 600
Oy      601 TAGGATTTGGAATGATGCTGACCTGCTGGAAGCTTTTCCATGACTATGAGATGATAC 660
Db      601 TAGGATTTGGAATGATGCTGACCTGCTGGAAGCTTTTCCATGACTATGAGATGATAC 660
Oy      661 AAGATGTTGAGAGATCTTCAGATTTAGCCAAACCAAAATTTGCTGAGATTAATAATGGG 720
Db      661 AAGATGTTGAGAGATCTTCAGATTTAGCCAAACCAAAATTTGCTGAGATTAATAATGGG 720
Oy      721 GCCTTGCTCCTCACTACTAGACACTTGTGTGCAAGAGCTCTGAAAGCCAAACAGATAC 780
Db      721 GCCTTGCTCCTCACTACTAGACACTTGTGTGCAAGAGCTCTGAAAGCCAAACAGATAC 780
Oy      781 GGCCTTGGAACCTGGAGGTTTATCTCTGTCTCAAGCAGACAGTAAATACGAGACAGG 840
Db      781 GGCCTTGGAACCTGGAGGTTTATCTCTGTCTCAAGCAGACAGTAAATACGAGACAGG 840
Oy      841 ATGCTTATGCTTACGATGCGATCTTACAAAGGTTCTTAAGGACCTCTGATGCTGCAGTG 900
Db      841 ATGCTTATGCTTACGATGCGATCTTACAAAGGTTCTTAAGGACCTCTGATGCTGCAGTG 900
Oy      901 GCTCTAATACGTGAAGAGAGAGAGCTTAAAGGTTAGCCCTTAACCCCAAGGTTAGCATCA 960
Db      901 GCTCTAATACGTGAAGAGAGAGAGCTTAAAGGTTAGCCCTTAACCCCAAGGTTAGCATCA 960
Oy      961 ATGATATGATACACCTAATCTAGTCAAGTATGCAATCTTGTGAATTTGATCTAGT 1020
Db      961 ATGATATGATACACCTAATCTAGTCAAGTATGCAATCTTGTGAATTTGATCTAGT 1020
Oy      1021 TCTGGTCCCTTTAAACCGTCCAGAAACTAG 1049
Db      1021 TCTGGTCCCTTTAAACCGTCCAGAAACTAG 1049

```

RESULT 2
; Sequence 1, Application US/09896186C

```

; GENERAL INFORMATION:
; APPLICANT: Joshua Z. Levin
; APPLICANT: Ken Phillips
; APPLICANT: Greg Budziszewski
; APPLICANT: Fred Meins
; APPLICANT: Zhenya Glazov
; TITLE OF INVENTION: Methods of Controlling Gene Expression
; FILE REFERENCE: 3148USNP
; CURRENT APPLICATION NUMBER: US/09/896,186C
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/222,202
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1

```

```

; LENGTH: 942
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-896-186C-1

```

```

Query Match      79.2%; Score 831; DB 5; Length 942;
Best Local Similarity 96.1%; Pred. No. 2,2e-219;
Matches 852; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

```

```

Oy      42 ATGTCATGCTCAAAATGATGATGAGCGACGCTTTTACAGAGAGAGAGCTTCCTATCGAC 101
Db      1 ATGTCATGCTCAAAATGATGATGAGCGACGCTTTTACAGAGAGAGAGCTTCCTATCGAC 60
Oy      102 GCCATCGAAGCTTCTTACAAATTTCCCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCCG 161
Db      61 GCCATCGAAGCTTCTTACAAATTTCCCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCCG 120
Oy      162 ACCGTACAAAGCTACAACTCCGCTCCATGCGCCACGAGAGAGATCCAAATTAATCCCAAT 221
Db      121 ACCGTACAAAGCTACAACTCCGCTCCATGCGCCACGAGAGAGATCCAAATTAATCCCAAT 180
Oy      222 AATATCCGTGCGCAATTTGCTGCTGCTCATCACTCTTCTTACATCTTATTAAGGATTTCT 281
Db      181 AATATCCGTGCGCAATTTGCTGCTGCTCATCACTCTTCTTACATCTTATTAAGGATTTCT 240
Oy      282 CTCTCCCGTTGCCGAGCTAGAAATTTTCCAGCAATGAGGTTGTGTAGATTTGTAT 341
Db      241 CTCTCCCGTTGCCGAGCTAGAAATTTTCCAGCAATGAGGTTGTGTAGATTTGTAT 300
Oy      342 AGCAAGACTGCTACGAGGTTGATTAAGCAGCAATGACACTTATTAATTTCTGTATAC 401
Db      301 AGCAAGACTGCTACGAGGTTGATTAAGCAGCAATGACACTTATTAATTTCTGTATAC 360
Oy      402 AAGAGAGATGAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 461
Db      361 AAGAGAGATGAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
Oy      462 AGAAAAAGTGTCTCCCGGGAAGGTTGCGACTGTCAGATATGATGATGATGATGATGAT 521
Db      421 AGAAAAAGTGTCTCCCGGGAAGGTTGCGACTGTCAGATATGATGATGATGATGATGAT 480
Oy      522 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 581
Db      481 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
Oy      582 GATTCACACTGTTAAAGGTAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 641
Db      541 GATTCACACTGTTAAAGGTAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
Oy      642 GACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 701
Db      601 GACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Oy      702 GGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 761
Db      661 GGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
Oy      762 CTGAAGCCAAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 821
Db      721 CTGAAGCCAAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
Oy      822 TTACATAGCGAGCAACGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 881
Db      781 TTACATAGCGAGCAACGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
Oy      882 CTTCCTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 928
Db      841 AAAAACCACTTCTACACTCAACGACCTTGAAGCAAAATCTCACA 887

```

RESULT 3
US-10-424-599-84537
; Sequence 84537, Application US/10424599

GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Tongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 84537
LENGTH: 1253
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_47352C.1
US-10-424-599-84537

Query Match 26.3%; Score 276.2; DB 8; Length 1253;
Best Local Similarity 66.6%; Pred. No. 8.9e-66;
Matches 395; Conservative 0; Mismatches 198; Indels 0; Gaps 0;

QY 300 AGGAATTTCCAGCAATGAGTTGGTGTAGATTTTGTATGCAAGACTGCTACTGAG 359
DB 276 ATGACATTTGCCCTGTAATGAGTTAGTGTCAAAATTTCTTATACAGAGACTTTTGATGCT 335
QY 360 GTTGATAGCGAGCAATGCACTTATTAAGTTCTTATACCAAGAGATGAATCTGGA 419
DB 336 GTGAGAGAACTGCACAAAGCTCTTACAAATTTCTCCAGAAAGACCGAGATGATG 395
QY 420 ATAGCTTTTGTGGCTGGATTAATGATGAGTGAACCAAGTTTAAAGAGTTCTGCCG 479
DB 396 CAACACTGCAATTTGATTTGACATTTGAGTGAACCCACTTCAGAAAAGTGTCTCC 455
QY 480 GGAAGGTTCCGACTGTCAGATATGTGTAGATTAATTAATTTGATGATTTATGCAAT 539
DB 456 GGAAGGTTCCGACTGTCAGATATGTGTAGATTAATTAATTTGATGATTTATGCAAT 515
QY 540 TTTTCATTTCTGTAATCCCTCAAGTCCTCAACATTTTATGAAGTTCAACACTGTGAAG 599
DB 516 ATTCATTTTCTGTAATCCCTCAAGTCCTCAACATTTTATGAAGTTCAACACTGTGAAG 575
QY 600 GTAGATTTGGAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 659
DB 576 GTTGGAGCTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 635
QY 660 AAAGATTTGAGATTTTCAAGATTTAGCAACCAAAATTTGTTGAGATTAATAAATG 719
DB 636 AAAGATTTGAGATTTTCAAGATTTAGCAACCAAAATTTGTTGAGATTAATAAATG 695
QY 720 GGCCTTGCTCCTCACTACTGAGACACTGTTTGCAGAAAGCTCCGAAAGCAAGATC 779
DB 696 GGCCTTGCTCCTCACTACTGAGACACTGTTTGCAGAAAGCTCCGAAAGCAAGATC 755
QY 780 AGCCTTGGAAGTGGAGTGGATTTTATCTCTGTCAAGAGCAGATTAATAGCGACAAG 839
DB 756 AGCCTTGGAAGTGGAGTGGATTTTATCTCTGTGTCAAGAGCAGATTAATAGCGACA 815
QY 840 GATGCTTATGCTTCAATGATGATTTTCAAGAGTCTTAAAGCACTTCTGATGC 892
DB 816 GATGCTTATGCTTCTTCTGTGTCTTATCAAGGCAATTAAGATCTCCGAGAGC 868

RESULT 4
US-10-389-048-9441
Sequence 9441, Application US/10389048
GENERAL INFORMATION:
APPLICANT: Hayukela, Ilkka
APPLICANT: Shenk, Michael A.
TITLE OF INVENTION: Polynucleotides, Materials Incorporating
FILE REFERENCE: 11000.10410U1
FILE REFERENCE: 11000.10410U1

CURRENT APPLICATION NUMBER: US/10/389,048
CURRENT FILING DATE: 2003-03-13
NUMBER OF SEQ ID NOS: 25129
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 9441
LENGTH: 672
TYPE: DNA
ORGANISM: Eucalyptus grandis
US-10-389-048-9441

Query Match 17.3%; Score 181.2; DB 8; Length 672;
Best Local Similarity 60.9%; Pred. No. 1.4e-39;
Matches 313; Conservative 0; Mismatches 198; Indels 3; Gaps 1;

QY 382 TTATTAAGTCTTGATACCAAGAGATGAATGTGAAATGCTTTTGTGCTTGATTA 441
DB 1 TTTTGAAGAAATATCAATGCCATGAAGATGAATGATCAATGATTAATTTGGCTTGACA 60
QY 442 TTGAGTGAGACCAAGTTTGAAGAGGTGTTCTCCGGGAGGTTCCGACTGTCAGA 501
DB 61 TTGAGTATAGGCCCAATTTTCAAAAAGGTGTTTCAACCTCGGAAGAGCATTAATGAGA 120
QY 502 TATGTATAGATTAATTAATTTATGATGATTTATGATTAATTTTCAATTTGTTGTTCC 561
DB 121 TTGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
QY 562 GTTCCCAACATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 621
DB 181 GTTGAAGCTTTATTTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 622 ACTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 681
DB 241 ATGCTAGAGAGGCTGTTAAGATTTCAATGATGATGATGATGATGATGATGATGATG 300
QY 682 ATTTAGCCCAACCAAAATTTG--GTGAGATTAATAAATGAGGCTTCCCTCACTAC 738
DB 301 AACTGTTGATTAATAAAGCTTCCATGCTGTTAATGATGATGATGATGATGATGATG 360
QY 739 AGACATGTTTGAAGAGTCTGGAAGCCCAACAGATGATGATGATGATGATGATG 798
DB 361 AGTGTCTTGAATTAAGAGTCTGGAAGCCCAACAGATGATGATGATGATGATGATG 420
QY 799 TTTATCCCTGCTCAAGCAGATTAATTAAGCAGCAGCAGCAGCAGCAGCAGCAGC 858
DB 421 CTCTTTTATTAATAAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 480
QY 859 ATCTTTACAGGTTCTTAAGAGCTTCTGATGC 892
DB 481 ATCTATATCAGCTCTTAAGAGCTGCGCAGATGC 514

RESULT 5
US-09-531-113-12947
Sequence 12947, Application US/09531113
GENERAL INFORMATION:
APPLICANT: Byrum, Joseph R.
APPLICANT: Heck, Gregory R.
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(15761)B
CURRENT APPLICATION NUMBER: US/09/531,113
CURRENT FILING DATE: 2000-03-22
NUMBER OF SEQ ID NOS: 48629
SEQ ID NO 12947
LENGTH: 579
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: jC-gmfl02220143602a1
US-09-531-113-12947
Query Match 13.6%; Score 142.4; DB 5; Length 579;

Best Local Similarity 64.9%; Pred. No. 7.5e-29;
Matches 226; Conservative 0; Mismatches 121; Indels 1; Gaps 1;

```
OY 300 AGGATTTTCCAGCAATGAGTTGGTGTAGATTTTGTATAGCAGACCTCTACTGAG 359
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 228 ATGACATTCCTGCTATATGATGATTTAGTGTCAAAATTTCTTATAGCAGACCTTTGATGCT 287
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 360 GTTGTAAAGCGACCAATCAGCTTATTAAGTCTTGATACCAAGAGATGATATCTGGA 419
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 288 GTAGAGAAAGCTGCAAAACAGCTTACAAATTTCCAAAGAAAACGACGACATGATG 347
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 420 ATAGCTTTTGGCTTGTGATTTAGTGAGACCAAGTTTATAGAAAAGTTTCTCCCG 479
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 348 CAAACTGCAATTTGATTTGACATTTAGTGGAACCCACCTTCA-AAAAGTCTTCTCC 406
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 480 GGAAGGTTGCGACATGTCAGATATGTGATAGTAAATATTTGATGATTTGATAT 539
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 407 GGAAGGTTGCGACATGTCAGATATGTGATAGTAAATATTTGATGATTTGATAT 466
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 540 TTTCAATTCGTATCCCTCAAAAGTCCCAACATCTTATTTGAAGATTCACACTTGAAG 599
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 467 ATTCAATTCGGAATCCCTCAAAATTTTACAGCTTTTGTGTTGAAGATCCACACTTTGAAG 526
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 600 GTAGCTATTGCAATGATGAGTCTGTGAAAGCTTTTCCATGACTAT 647
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 537 GTTGAGCTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 574
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 6

US-09-531-113-12947
; Sequence 12947, Application US/09531113

; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Heck, Gregory R.
; APPLICANT: La Rosa, Thomas J.
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-21(15761)B
; CURRENT APPLICATION NUMBER: US/09/531,113
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 48629
; SEQ ID NO 12947
; LENGTH: 579
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: jC-gmfl02220143d02a1
US-09-531-113-12947

Query Match 13.6%; Score 142.4; DB 6; Length 579;
Best Local Similarity 64.9%; Pred. No. 7.5e-29;

Matches 226; Conservative 0; Mismatches 121; Indels 1; Gaps 1;

```
OY 300 AGGATTTTCCAGCAATGAGTTGGTGTAGATTTTGTATAGCAGACCTCTACTGAG 359
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 228 ATGACATTCCTGCTATATGATGATTTAGTGTCAAAATTTCTTATAGCAGACCTTTGATGCT 287
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 360 GTTGTAAAGCGACCAATCAGCTTATTAAGTCTTGATACCAAGAGATGATATCTGGA 419
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 288 GTAGAGAAAGCTGCAAAACAGCTTACAAATTTCCAAAGAAAACGACGACATGATG 347
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 420 ATAGCTTTTGGCTTGTGATTTAGTGAGACCAAGTTTATAGAAAAGTTTCTCCCG 479
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 348 CAAACTGCAATTTGATTTGACATTTAGTGGAACCCACCTTCA-AAAAGTCTTCTCC 406
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 480 GGAAGGTTGCGACATGTCAGATATGTGATAGTAAATATTTGATGATTTGATAT 539
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 407 GGAAGGTTGCGACATGTCAGATATGTGATAGTAAATATTTGATGATTTGATAT 466
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 540 TTTCAATTCGTATCCCTCAAAAGTCCCAACATCTTATTTGAAGATTCACACTTGAAG 599
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 467 ATTCAATTCGGAATCCCTCAAAATTTTACAGCTTTTGTGTTGAAGATCCACACTTTGAAG 526
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 600 GTAGCTATTGCAATGATGAGTCTGTGAAAGCTTTTCCATGACTAT 647
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

DB 527 GTTGAGCTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 574
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 7

US-09-531-113-705
; Sequence 705, Application US/09531113

; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Heck, Gregory R.
; APPLICANT: La Rosa, Thomas J.
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-21(15761)B
; CURRENT APPLICATION NUMBER: US/09/531,113
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 48629
; SEQ ID NO 705
; LENGTH: 278
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700909421H1
US-09-531-113-705

Query Match 11.9%; Score 124.6; DB 5; Length 278;
Best Local Similarity 68.5%; Pred. No. 5e-24;
Matches 172; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

```
OY 642 GACTATGAGTATGATATCAAGATGTTGAGATCTTTACGATTTAGCCACCAAAAAT 701
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4 GATATATACATATCTGTTAAAGTGTGACGATCTTTTTCATGATCAAAAAGCTT 63
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 702 GGTGAGATAAAAATGGGCGCTGCTGCTACAGTACGACACTGTTTGAAGAGCTC 761
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 64 GGTGAGATCAATGATGAGGCTCTTGATCTTTGACTGAAAACCTTATCAAAACAGCTT 123
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 762 CTGAAGCCCAAGATCAGGCTTGGGAACGAGATTTTATCTCTGTCAAAAGCAG 821
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 124 AAAAGCCCAAAATAAGCTGGGAATTTGGAGGCTCTGTTTGTCAAAAGGACAA 183
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 822 TTACATATGACGACGATGCTTATGCTTATGATGATCTTTCAAGTCTTGAAGAC 881
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 184 CTAGATATGCTGCAACAGATCTTTGCTTGTGTGCTTATACGAGCATTAAGAT 243
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 882 CTCTCGATGC 892
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 244 CTCCCGGAGCC 254
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 8

US-09-531-113-705
; Sequence 705, Application US/09531113

; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Heck, Gregory R.
; APPLICANT: La Rosa, Thomas J.
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-21(15761)B
; CURRENT APPLICATION NUMBER: US/09/531,113
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 48629
; SEQ ID NO 705
; LENGTH: 278
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 700909421H1
US-09-531-113-705

Query Match 11.9%; Score 124.6; DB 6; Length 278;
Best Local Similarity 68.5%; Pred. No. 5e-24;
Matches 172; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

```
QY 642 GACTATGAGATTAGTATCAAGATCTTGAGATCTTTCAGATTAGCCACCAAAATTT 701
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 4 GATTATACATATCTGTTAAAGGTGTGACGATCTTTTTCATGCTAAATCAAAAGCTT 63
QY 702 GGTGAGATATAAAATGGGGCTTGCCCTCACTAATCAGACACTGTTTGGAAAGAGCTC 761
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 64 GGTGAGATCATATAGTGGGCTTGCCATCTTGTGACTGTAATAAATCTTCAAAACAGCTT 123
QY 762 CTGAAGCCAAACAGATAGCGCTTGGGACTGGAGTTTATCTCTGTCAAGACGAG 821
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 124 AAAAAGCCCAAAATTAAGACTGGGAATTTGGGAGCGCTCTGTTTGTCAAGAGCAA 183
QY 822 TTACAAATACGACAGAGATGCTTATGCTTACAGGATCTTTACAGGTTCTTAAGAC 881
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 184 CTAGAGTATGCTGCACACAGATGCTTTGCTTGTGCTTATACAGCGCATTAAGAT 243
QY 882 CTTCCTGATGC 892
    ||||| | | | | |
Db 244 CTCCCGAGCG 254
```

RESULT 9

```
US-09-615-606A-75994
; Sequence 75994, Application US/09615606A
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; APPLICANT: Buehler, Robert E.
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Coombs, Brian E.
; APPLICANT: Heck, Gregory R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Nelson, Donald E.
; APPLICANT: Shukla, Hridayabhiranjan
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(1544)C
; CURRENT APPLICATION NUMBER: US/09/615,606A
; CURRENT FILING DATE: 2000-07-13
; NUMBER OF SEQ ID NOS: 91663
; SEQ ID NO 75994
; LENGTH: 278
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700909421H1
US-09-615-606A-75994
```

Query Match 11.9%; Score 124.6; DB 6; Length 278;

Best Local Similarity 68.5%; Pred. No. 5e-24; Mismatches 79; Indels 0; Gaps 0;

```
Matches 172; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
QY 642 GACTATGAGATTAGTATCAAGATCTTGAGATCTTTCAGATTAGCCACCAAAATTT 701
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 4 GATTATACATATCTGTTAAAGGTGTGACGATCTTTTTCATGCTAAATCAAAAGCTT 63
QY 702 GGTGAGATATAAAATGGGGCTTGCCCTCACTAATCAGACACTGTTTGGAAAGAGCTC 761
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 64 GGTGAGATCATATAGTGGGCTTGCCATCTTGTGACTGTAATAAATCTTCAAAACAGCTT 123
QY 762 CTGAAGCCAAACAGATAGCGCTTGGGACTGGAGTTTATCTCTGTCAAGACGAG 821
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 124 AAAAAGCCCAAAATTAAGACTGGGAATTTGGGAGCGCTCTGTTTGTCAAGAGCAA 183
QY 822 TTACAAATACGACAGAGATGCTTATGCTTACAGGATCTTTACAGGTTCTTAAGAC 881
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 184 CTAGAGTATGCTGCACACAGATGCTTTGCTTGTGCTTATACAGCGCATTAAGAT 243
QY 882 CTTCCTGATGC 892
    ||||| | | | | |
Db 244 CTCCCGAGCG 254
```

RESULT 10

```
US-09-724-676-40395
; Sequence 40395, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Comugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Comugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 40395
; LENGTH: 3255
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-40395
```

Query Match 10.5%; Score 111; DB 6; Length 3255;

Best Local Similarity 54.1%; Pred. No. 6.5e-20; Mismatches 225; Indels 6; Gaps 2;

```
Matches 272; Conservative 0; Mismatches 225; Indels 6; Gaps 2;
QY 396 GATACCAAGAGAGATGAATCTGGAATAGCTTTGTGGCTTGATATGAGTGAGACCA 455
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 433 GATTATAGCATGAGTCTATACAGATGAGGAGTGTGGGATTTGACATGAGTGAGCCACCA 492
QY 456 AGTTTACAAAAGGTGTTCTCCCGGGGAAGGTGGGACTGTCCAGATATGTATAGTACT 515
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 493 TTATACAAATAGAGGGAACCT--TGCCAAAGTTGCACTAATTCAGTTGTTGCTTGAG 549
QY 516 AATTATATGATGATGATGATATTTT---CATTCGGATTCCTCAAAAGTCCCAACAT 572
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 550 AGCAATATGATCTTCTTCCACGTTCTTCCATGATGATTTTCCCAAGGATTAATAATG 609
QY 573 CTATATGAAGATTACCACTTGTAAAGTAGTATGGAATGATGATGATCTGTGAG 632
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 610 TTGCTTGAATAATTAAGCAAGTTAAAGGAGGAGTAGAATTAAGAGAGATCAGTGGA 669
QY 633 CTTTCCATGATATAGAGTTAGTATCAAAAGATTTGGAGATCTTCCAGATTTAGCAAC 692
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 670 CTTCTACGAGCTTGTGATATCAAAATTTGAAGATTTTGGAGTTTACAGATGTTCCCAT 729
QY 693 CAAAAAATTTGGGAGATTAATAATGGGCTTGCTCCTCACTAGACACTGTGTTGC 752
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 730 AAAAAGCTGAATGTACAGAGACCTGAGACCTTAACAGTCTGTAAACACCTTTAGT 789
QY 753 AAAGAGCTCTGTAGAACCAAAAGCAATCAAGGCTTGGAACTGGAGCTTTATCTGTCA 812
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 790 AAAGAGCTCTGTAGAACCAAGTCAATCCGCTGTAGCAATTTGAGTAATTTCTCTCACT 849
QY 813 AAGCAGCATTTACATAGCAGCAACGATGCTTATGCTTCAATGAGCATTTACAAAGTT 872
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 850 GAGGACCGAAMAACTATAGCAGCCACTGATGCTTATGCTTATTTATTTACCGAAT 909
QY 873 CTTAGGACCTTCCTGAGCTGT 895
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 910 TTAGAGATTTTGAATGATCTGT 932
```

RESULT 11

```
US-09-724-676A-40395
; Sequence 40395, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Comugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Comugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 40395
; LENGTH: 3255
; TYPE: DNA
```

ORGANISM: Homo sapiens
US-09-724-676A-40395

Query Match 10.6%; Score 111; DB 6; Length 3255;
Best Local Similarity 54.1%; Pred. No. 6, 5e-20;
Matches 272; Conservative 0; Mismatches 225; Indels 6; Gaps 2;

QY 396 GATACCAAGAGATGATCTGTAATAGCTTTTGTGGCTTGATATGATGAGACCA 455
DB 433 GATATTAGCATGATGCTTATCATGATGGGATGGTGGATTTGACATGAGAGCCACCA 492
QY 456 AGTTTGAAGAAAGGTGTTCCCGGGGAAGGTTGCGACTGTCACATATGATAGTACT 515
DB 493 TTATACAAATAGAGGGAACCT---TGCCAAAGTTGCACTAATTCAGTTGTGTCTTGAG 549
QY 516 AATTATGTGATGTTATGATATTTT---CATTCGTATCCCTCAAGTCCCAACAT 572
DB 550 AGCAATGTACTGTGTCACGTTTCTTCATGTCAGTTTTCGCCAGGATTTAAATATG 609
QY 573 CTATATTGAAGATCAACACTTGTAAAGGTATTTGAATGTAGTACTCTGTGAAG 632
DB 610 TTGCTTGAATAAAGCAGTTAAAGGCGAGGTGTGAGAAATGAAGAGATCAGTGA 669
QY 633 CTTTTCAGTACTATGAGATAGTATCAAGAATGTTGAGAGATCTTCAATTTAGCCAC 692
DB 670 CTTTACGCTGACTTTGATATCAAAATGAAGAAATTTGTGAGTTGACAGATGTTGCCAAT 729
QY 693 CAAAAAATGTGTGAGATTAATAAATGGGCTTGCCCTCACTAAGTGAAGACTGTTTGC 752
DB 730 AAAAGCTCAATATGACAGAGACCTGAGCCTTAACAGTCTGTTAAACACTCTTAGGT 789
QY 753 AAAGACTCTCTGAAGCAAAACAGATCAGGCTTGGAGACTGGAGTTTATCCTCTGTCA 812
DB 790 AAACAGCTCTGAAAGACAGCTATCCGCTGACAAATTTGGAATTTCCCTCAGCT 849
QY 813 AAGCAGCAGTTACATACGCAACGAGATGCTTATGCTTCAATGATCTTTACAAAGTT 872
DB 850 GAGGACCGAAACTGATGACGCCAGTATGCTTATGCTGTTTATTTATTTACCGAAAT 909
QY 873 CTTAAGGACCTTCCTGATGCTGT 895
DB 910 TTAGAGATTTTGGATGATGATCTGT 932

RESULT 12
US-09-896-186C-17
; Sequence 17, Application US/09896186C
; GENERAL INFORMATION:
; APPLICANT: Joshua Z. Levin
; APPLICANT: Ken Phillips
; APPLICANT: Greg Budziszewski
; APPLICANT: Fred Meins
; APPLICANT: Zhenya Glazov
; TITLE OF INVENTION: Methods of Controlling Gene Expression
; FILE REFERENCE: 31481USNP
; CURRENT APPLICATION NUMBER: US/09/896,186C
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/222,202
; PRIOR FILING DATE: 2000-08-01
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 4299
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-896-186C-17

Query Match 10.6%; Score 111; DB 5; Length 4299;
Best Local Similarity 54.1%; Pred. No. 7, 2e-20;
Matches 272; Conservative 0; Mismatches 225; Indels 6; Gaps 2;
QY 396 GATACCAAGAGATGATCTGTAATAGCTTTTGTGGCTTGATATGATGAGACCA 455
DB 433 GATATTAGCATGATGCTTATCATGATGGGATGGTGGATTTGACATGAGAGCCACCA 492
QY 456 AGTTTGAAGAAAGGTGTTCCCGGGGAAGGTTGCGACTGTCACATATGATAGTACT 515
DB 493 TTATACAAATAGAGGGAACCT---TGCCAAAGTTGCACTAATTCAGTTGTGTCTTGAG 549
QY 516 AATTATGTGATGTTATGATATTTT---CATTCGTATCCCTCAAGTCCCAACAT 572
DB 550 AGCAATGTACTGTGTCACGTTTCTTCATGTCAGTTTTCGCCAGGATTTAAATATG 609
QY 573 CTATATTGAAGATCAACACTTGTAAAGGTATTTGAATGTAGTACTCTGTGAAG 632
DB 610 TTGCTTGAATAAAGCAGTTAAAGGCGAGGTGTGAGAAATGAAGAGATCAGTGA 669
QY 633 CTTTTCAGTACTATGAGATAGTATCAAGAATGTTGAGAGATCTTCAATTTAGCCAC 692
DB 670 CTTTACGCTGACTTTGATATCAAAATGAAGAAATTTGTGAGTTGACAGATGTTGCCAAT 729
QY 693 CAAAAAATGTGTGAGATTAATAAATGGGCTTGCCCTCACTAAGTGAAGACTGTTTGC 752
DB 730 AAAAGCTCAATATGACAGAGACCTGAGCCTTAACAGTCTGTTAAACACTCTTAGGT 789
QY 753 AAAGACTCTCTGAAGCAAAACAGATCAGGCTTGGAGACTGGAGTTTATCCTCTGTCA 812
DB 790 AAACAGCTCTGAAAGACAGCTATCCGCTGACAAATTTGGAATTTCCCTCAGCT 849
QY 813 AAGCAGCAGTTACATACGCAACGAGATGCTTATGCTTCAATGATCTTTACAAAGTT 872
DB 850 GAGGACCGAAACTGATGACGCCAGTATGCTTATGCTGTTTATTTATTTACCGAAAT 909
QY 873 CTTAAGGACCTTCCTGATGCTGT 895
DB 910 TTAGAGATTTTGGATGATGATCTGT 932

DB 202 GATATTAGCATGATCTATCAGATGGGATGTGGTGGATTTGACATGAGAGCCACCA 261
QY 456 AGTTTGAAGAAAGGTGTTCCCGGGGAAGGTTGCGACTGTCACATATGATGATAGTACT 515
DB 262 TTATACAAATAGAGGGAACCT---TGCCAAAGTTGCACTAATTCAGTTGTGTCTTGAG 318
QY 516 AATTATGTGATGTTATGATATTTT---CATTCGTATCCCTCAAGTCCCAACAT 572
DB 319 AGCAATGTACTGTGTCACGTTTCTTCATGTCAGTTTTCGCCAGGATTTAAATATG 378
QY 573 CTATATTGAAGATCAACACTTGTAAAGGTATTTGAATGTAGTACTCTGTGAAG 632
DB 379 TTGCTTGAATAAAGCAGTTAAAGGCGAGGTGTGAGAAATGAAGAGATCAGTGA 438
QY 633 CTTTTCAGTACTATGATGATATCAAGAAGTGTGAGAGATCTTCAATTTAGCCAC 692
DB 439 CTTTACGCTGACTTATATCAATTTGAAGAAATTTGTGAGAGTGTACATATGTTGCCAAT 498
QY 693 CAAAAAATGTGTGAGATTAATAAATGGGCTTGCCCTCACTAAGTGAAGACTGTTTGC 752
DB 499 AAACAGCTCTGAAAGACAGACCTGAGCCTTAACAGTCTGTTAAACACTCTTAGGT 558
QY 753 AAAGACTCTCTGAAGCAAAACAGATCAGGCTTGGAGACTGGAGTTTATCCTCTGTCA 812
DB 559 AAACAGCTCTGAAAGACAGACCTGAGCCTTAACAGTCTGTTAAACACTCTTAGGT 618
QY 813 AAGCAGCAGTTACATACGCAACGAGATGCTTATGCTTCAATGATGATCTTTACAAAGTT 872
DB 619 GAGGACCAAGAACTGATGACGCCAGTATGCTTATGCTGTTTATTTATTTACCGAAAT 678
QY 873 CTTAAGGACCTTCCTGATGCTGT 895
DB 679 TTAGAGATTTTGGATGATGATCTGT 701

RESULT 13
US-09-949-001-3
; Sequence 3, Application US/09949001
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH OSTEOPOROSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL000788
; CURRENT APPLICATION NUMBER: US/09/949,001
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/231,323
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 848
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 5208
; TYPE: DNA
; ORGANISM: Human
US-09-949-001-3

Query Match 10.6%; Score 111; DB 6; Length 5208;
Best Local Similarity 54.1%; Pred. No. 7, 6e-20;
Matches 272; Conservative 0; Mismatches 225; Indels 6; Gaps 2;

QY 396 GATACCAAGAGATGATCTGTAATAGCTTTTGTGGCTTGATATGATGAGAGACCA 455
DB 433 GATATTAGCATGATGCTTATCATGATGGGATGGTGGATTTGACATGAGAGCCACCA 492
QY 456 AGTTTGAAGAAAGGTGTTCCCGGGGAAGGTTGCGACTGTCACATATGATGATAGTACT 515
DB 493 TTATACAAATAGAGGGAACCT---TGCCAAAGTTGCACTAATTCAGTTGTGTCTTGAG 549
QY 516 AATTATGTGATGTTATGATATTTT---CATTCGTATCCCTCAAGTCCCAACAT 572
DB 550 AGCAATGTACTGTGTCACGTTTCTTCATGTCAGTTTTCGCCAGGATTTAAATATG 609
QY 573 CTATATTGAAGATCAACACTTGTAAAGGTATTTGAATGTAGTACTCTGTGAAG 632
DB 610 TTGCTTGAATAAAGCAGTTAAAGGCGAGGTGTGAGAAATGAAGAGATCAGTGA 669
QY 633 CTTTTCAGTACTATGATGATATCAAGAAGTGTGAGAGATCTTCAATTTAGCCAC 692
DB 439 CTTTACGCTGACTTATATCAATTTGAAGAAATTTGTGAGAGTGTACATATGTTGCCAAT 498
QY 693 CAAAAAATGTGTGAGATTAATAAATGGGCTTGCCCTCACTAAGTGAAGACTGTTTGC 752
DB 499 AAACAGCTCTGAAAGACAGACCTGAGCCTTAACAGTCTGTTAAACACTCTTAGGT 558
QY 753 AAAGACTCTCTGAAGCAAAACAGATCAGGCTTGGAGACTGGAGTTTATCCTCTGTCA 812
DB 559 AAACAGCTCTGAAAGACAGACCTGAGCCTTAACAGTCTGTTAAACACTCTTAGGT 618
QY 813 AAGCAGCAGTTACATACGCAACGAGATGCTTATGCTTCAATGATGATCTTTACAAAGTT 872
DB 619 GAGGACCAAGAACTGATGACGCCAGTATGCTTATGCTGTTTATTTATTTACCGAAAT 678
QY 873 CTTAAGGACCTTCCTGATGCTGT 895
DB 679 TTAGAGATTTTGGATGATGATCTGT 701


```
Db 611 TTGCTGAAAATTAAGCAGTTAAAAAGCAGGTGTAGAAATTGAAGAGATCAGTGAAA 670
QY 633 CTTTCCATGACTATGAGTAGTATCAAAAGATGTGAGATCTTTCAGATTAGCCAA 692
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 671 CTTCTACGTGACTTTGATATCAAAATTGAAGAAATTTGTGAGCTTGACAGATGTGCCAA 730
QY 693 CAAAAAATTTGGTGAGATAAAAAATGGGCGCTTGCTCACTAAGACACTTGTTC 752
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 731 AAAAGCTGAAATGCACAGAGACCTGAGCCCTTAACAGTCTGTTAAACACCTTAGGT 790
QY 753 AAAAGCTGCTGAAGCCAAACAGAAATCAGGCTTGGGAACCTGGAGTTTATCCTGTCA 812
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 791 AAACAGCTCCTGAAAGACAGTCTATCCGCTGAGCAATTGAGTAAATTTCTCTCACP 850
QY 813 AAGCAGCACTTACATATACGACAGGATGCTTATGCTTCATGCGCATCTTACAAAGTT 872
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 851 GAGGACCAAGAACTGTATGACAGCCACTGATGCTTATGCTGTTTATATTATTTACGAAAT 910
QY 873 CTTAAGGACCTTCTGATGCTGT 895
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 911 TTAGAGATTTTGGATGATACTGT 933
```

Search completed: June 14, 2003, 09:27:35
Job time : 920 secs

PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135153.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142877.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144086.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159395.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.

PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 100.0%; Score 1491; DB 21; Length 288;
 Best Local Similarity 100.0%; Pred. No. 5,6e-151;
 Matches 288; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSNMWIDDAFTTEELLALDAIEASYNFSRSSSSSSSAAPVQATTSVHGHEEDPNQIPN 60
 DB 1 MSSNMWIDDAFTTEELLALDAIEASYNFSRSSSSSSSAAPVQATTSVHGHEEDPNQIPN 60
 QY 61 NNRQLPRSTSTSYKRFPLSRCRARNFPMRFGGRILYKTADEVDRAMQLIKVLD 120
 DB 61 NNRQLPRSTSTSYKRFPLSRCRARNFPMRFGGRILYKTADEVDRAMQLIKVLD 120
 QY 121 KRDESGIAFVGIDTEMPSFRKGYLPKGVATVQICVDSNYCDVMIIFHSGLPQSLQHLIE 180
 DB 121 KRDESGIAFVGIDTEMPSFRKGYLPKGVATVQICVDSNYCDVMIIFHSGLPQSLQHLIE 180
 QY 181 DSTLVKVGIGIDGDSVKLFHDYGVSIKDEVEDLSLANQKIGDKRWGLASLTETLVCKEL 240
 DB 181 DSTLVKVGIGIDGDSVKLFHDYGVSIKDEVEDLSLANQKIGDKRWGLASLTETLVCKEL 240
 QY 241 LKPNRIRLGNWFEYPLSKQQLQYATDAAYASMHLYKVLKDLPDVAVSGS 288
 DB 241 LKPNRIRLGNWFEYPLSKQQLQYATDAAYASMHLYKVLKDLPDVAVSGS 288

RESULT 2
 AAM50937 standard; Protein: 288 AA.
 AC AAM50937;
 DT 15-MAY-2002 (first entry)
 DE Arabidopsis RNase D related domain (3'-5' exonuclease domain).
 KW RNase D; enzyme; 3'-5' exonuclease; gene expression;
 OS gene silencing; transgenic plant.
 XX Arabidopsis thaliana.
 XX MO200210362-A2.
 PD 07-FEB-2002.
 PE 30-JUL-2001; 2001WO-EP08825.
 PR 01-AUG-2000; 2000US-222202P.
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.

PA (NOVS) NOVARTIS FORSCHUNGSSTIFTUNG ZWEIGNIEDERL.
 PI Levin JZ, Phillips KL, Budziszewski GJ, Meins F, Glazov EA;
 DR WPI; 2002-217119/27.
 PR N-PSDB; ABA91805.
 PT Regulating gene expression in plants for controlling gene silencing,
 PT comprises altering the transcription or translation of an endonuclease
 PT nucleotide sequence encoding a polypeptide comprising an exonuclease
 domain -
 Claim 1; Page 94-95; 102pp; English.

The present is that of an Arabidopsis thaliana 3'-5' exonuclease domain, specifically an RNase D related domain. The 3'-5' exonuclease domain consists of 3 sequence motifs termed ExoI, ExoII and ExoIII. It was identified on the basis of homology to the Arabidopsis 3'-5' exonuclease sequence given in AAM50927. The nucleotide (see ABA91795-807) and polypeptide (see AAM50927-39) sequences for 3'-5' exonuclease domains of the invention were identified using different screening methods and search algorithms. The invention encompasses the suppressing method or increase of gene silencing in plants. This is achieved by altering the expression in the plant cell of a nucleotide sequence encoding a polypeptide, such as the present sequence, having an exonuclease domain, preferably a 3'-5' exonuclease domain, and especially an RNase D related domain. Expression of the nucleotide sequence is altered by altering its transcription or translation. It is reduced e.g. by sense suppression, antisense suppression, homologous recombination, the use of ribozymes, dominant-negative mutants, aptamers, zinc finger proteins, double-stranded RNA, or insertional, point or deletion mutagenesis, and increased by overexpression. The methods are useful for altering or stabilizing the expression of a nucleotide sequence of interest in a plant cell or plant. Suppressing or decreasing expression of the nucleic acid molecule results in decreased levels of post transcriptional gene silencing and improved expression of genes of interest. The 3'-5' exonuclease domain polypeptides are also useful for identifying compounds capable of altering 3'-5' exonuclease activity.

Sequence 288 AA;
 Query Match 100.0%; Score 1491; DB 23; Length 288;
 Best Local Similarity 100.0%; Pred. No. 5,6e-151;
 Matches 288; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSNMWIDDAFTTEELLALDAIEASYNFSRSSSSSSSAAPVQATTSVHGHEEDPNQIPN 60
 DB 1 MSSNMWIDDAFTTEELLALDAIEASYNFSRSSSSSSSAAPVQATTSVHGHEEDPNQIPN 60
 QY 61 NNRQLPRSTSTSYKRFPLSRCRARNFPMRFGGRILYKTADEVDRAMQLIKVLD 120
 DB 61 NNRQLPRSTSTSYKRFPLSRCRARNFPMRFGGRILYKTADEVDRAMQLIKVLD 120
 QY 121 KRDESGIAFVGIDTEMPSFRKGYLPKGVATVQICVDSNYCDVMIIFHSGLPQSLQHLIE 180
 DB 121 KRDESGIAFVGIDTEMPSFRKGYLPKGVATVQICVDSNYCDVMIIFHSGLPQSLQHLIE 180
 QY 181 DSTLVKVGIGIDGDSVKLFHDYGVSIKDEVEDLSLANQKIGDKRWGLASLTETLVCKEL 240
 DB 181 DSTLVKVGIGIDGDSVKLFHDYGVSIKDEVEDLSLANQKIGDKRWGLASLTETLVCKEL 240
 QY 241 LKPNRIRLGNWFEYPLSKQQLQYATDAAYASMHLYKVLKDLPDVAVSGS 288
 DB 241 LKPNRIRLGNWFEYPLSKQQLQYATDAAYASMHLYKVLKDLPDVAVSGS 288

RESULT 3
 AAG52857 standard; Protein: 298 AA.
 ID AAG52857
 AC AAG52857;

```
XX 18-OCT-2000 (first entry)
DT Arabidopsis thaliana protein fragment SEQ ID NO: 67236.
XX
DE Protein identification; signal transduction pathway; metabolic pathway;
XX hybridization assay; genetic mapping; gene expression control; promoter;
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144532.
PR 20-JUL-1999; 99US-0144584.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145313.
PR 27-JUL-1999; 99US-0145318.
PR 27-JUL-1999; 99US-0145319.
PR 28-JUL-1999; 99US-0145351.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149175.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
```

```

PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158366.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match 100.0%; Score 1491; DB 21; Length 298;
Best Local Similarity 100.0%; Pred. No. 5.9e-151;
Matches 288; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MSSNNWDDATEEELLAIDAIEASYNSSSSSSSSSAAPVQATTSVHGHEEDNPQIPN 60
DB 11 MSSNNWDDATEEELLAIDAIEASYNSSSSSSSSSAAPVQATTSVHGHEEDNPQIPN 70
QY 61 NNRQLPRSTSSSYKRFPLSRARFPMARFGGRILYKTKATEVDKRAMOLIKVDT 120
DB 71 NNRQLPRSTSSSYKRFPLSRARFPMARFGGRILYKTKATEVDKRAMOLIKVDT 130
QY 121 KRDESGIAFVGLDIEMRPSFRKGYLPCKVATVQICVDNMYCDVMHIFHSGIPQSLQHLIE 180
DB 131 KRDESGIAFVGLDIEMRPSFRKGYLPCKVATVQICVDNMYCDVMHIFHSGIPQSLQHLIE 190
QY 181 DSTLVKKGIGIDGDSVKLFHGYGSIKDYEDLSLANCKIGDGKKMGSLATETLVCKEL 240
DB 191 DSTLVKKGIGIDGDSVKLFHGYGSIKDYEDLSLANCKIGDGKKMGSLATETLVCKEL 250
QY 241 LKPNRIRLGNWMEFYPLSKQQLQVYAATDAVASMHLKYVKLDPDAVSGS 288

```

```

DB 251 LKPNRIRLGNWMEFYPLSKQQLQVYAATDAVASMHLKYVKLDPDAVSGS 298
|||||
RESULT 4
AAM50927
AAM50927 standard; Protein; 313 AA.
XX
AC AAM50927;
XX
DT 15-MAY-2002 (first entry)
XX
DE Arabidopsis Rnase D related domain (3'-5' exonuclease domain).
XX
KW Rnase D; enzyme; 3'-5' exonuclease; gene expression; plant;
KM gene silencing; transgenic plant.
XX
OS Arabidopsis thaliana.
XX
FH Key Location/Qualifiers
FH Domain 129..287
FT /label= Exonuclease
FT Peptide 129..138
FT /label= Exoi
FT Peptide 186..201
FT /label= ExoII
FT Peptide 260..270
FT /label= ExoIII
XX
PN WO200210362-A2.
XX
PN 07-FEB-2002.
XX
PD 30-JUL-2001; 2001WO-EP06825.
XX
PR 01-AUG-2000; 2000US-222202P.
XX
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX (NOVS ) NOVARTIS FORSCHUNGSSTIFTUNG ZWEIGNIEDERL.
XX
PI Levin JZ, Phillips KL, Budziszewski GJ, Melns F, Glazov EA;
XX WPI; 2002-217119/27.
XX N-PSDB; ABA91795.
XX
PT Regulating gene expression in plants for controlling gene silencing,
PT comprises altering the transcription or translation of an endonuclease
PT nucleotide sequence encoding a polypeptide comprising an exonuclease
XX domain -
XX
PS Claim 1; Page 70-71; 102pp; English.
XX
CC The present is that of an Arabidopsis thaliana 3'-5' exonuclease
CC domain (GenPept accession CAB36851), specifically an Rnase D related
CC domain. The 3'-5' exonuclease domain consists of 3 sequence motifs
CC termed Exoi, Exoii and Exoiii. Nucleotide (see ABA91795-807) and
CC polypeptide (see AAM50927-39) sequences for 3'-5' exonuclease
CC domains were identified using different screening methods and search
CC algorithms. The invention encompasses the suppression or increase
CC of gene silencing in plants. This is achieved by altering the
CC expression in the plant cell of a nucleotide sequence encoding a
CC polypeptide, such as the present sequence, having an exonuclease
CC domain, preferably a 3'-5' exonuclease domain, and especially
CC an Rnase D related domain. Expression of the nucleotide sequence
CC is altered by altering its transcription or translation. It is
CC reduced e.g. by sense suppression, antisense suppression, homologous
CC recombination, the use of ribozymes, dominant-negative mutants,
CC aptamers, zinc finger proteins, double-stranded RNA, or insertional,
CC point or deletion mutagenesis, and increased by overexpression.
CC The methods are useful for altering or stabilizing the expression of
CC a nucleotide sequence of interest in a plant cell or plant.
CC Suppressing or decreasing expression of the nucleic acid molecule
CC results in decreased levels of post transcriptional gene silencing

```

CC and improved expression of genes of interest. The 3'-5' CC exonuclease domain polypeptides are also useful for identifying compounds capable of altering 3'-5' exonuclease activity.

XX Sequence 313 AA:

Query Match 96.5%; Score 1439.5; DB 23; Length 313;
Best Local Similarity 95.3%; Pred. No. 2.1e-145;
Matches 281; Conservative 1; Mismatches 4; Indels 9; Gaps 1;

QY 1 MSSSMWIDAFTEEEELATDALEASNFSSSSSSSSAAPTYQATTSVGHEDDPQIYN 60
DB 1 MSSSMWIDAFTEEEELATDALEASNFSSSSSSSSAAPTYQATTSVGHEDDPQIYN 60
QY 61 NTRROLPRSTSTSTSKRFPPLSRCRARNPAMRFGGRILYKTAPEVDRAMOLIKVLPD 120
DB 61 NTRROLPRSTSTSTSKRFPPLSRCRARNPAMRFGGRILYKTAPEVDRAMOLIKVLPD 120
QY 121 KRDESGIAFVGLDIEMRPSFRKGVLPKVAIVQICVDNVCVDMHIFHSGIFOSLOHLE 180
DB 121 KRDESGIAFVGLDIEMRPSFRKGVLPKVAIVQICVDNVCVDMHIFHSGIFOSLOHLE 180
QY 181 DSTLVKVGIGIDGDSYKLFHDYGSIKDYEDLSDLANOKIGGDKKGLASTETLVCKEL 240
DB 181 DSTLVKVGIGIDGDSYKLFHDYGSIKDYEDLSDLANOKIGGDKKGLASTETLVCKEL 240
QY 241 LKPNRIRLGNMEFFYPLSKOOLQYAATDAYASWHLXKV-----LKDLPDPAVS 286
DB 241 LKPNRIRLGNMEFFYPLSKOOLQYAATDAYASWHLXKVTTTKNHLTLNLEAKIS 295

RESULT 5
AAG24977
ID AAG24977 standard; Protein: 288 AA.

XX AAG24977;

XX DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 28856.

KW Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX PD 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134570.
PR 14-MAY-1999; 99US-0134768.
PR 18-MAY-1999; 99US-0134941.
PR 19-MAY-1999; 99US-0135124.
PR 20-MAY-1999; 99US-0135533.
PR 21-MAY-1999; 99US-0135629.
PR 24-MAY-1999; 99US-0136021.
PR 25-MAY-1999; 99US-0136392.
PR 27-MAY-1999; 99US-0136782.
PR 28-MAY-1999; 99US-0137222.
PR 01-JUN-1999; 99US-0137528.
PR 03-JUN-1999; 99US-0137502.
PR 04-JUN-1999; 99US-0137724.
PR 07-JUN-1999; 99US-0138094.
PR 08-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 10-JUN-1999; 99US-0139119.
PR 14-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 16-JUN-1999; 99US-0139454.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0147038.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.

PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 96.1%; Score 1433; DB 21; Length 288;
Best Local Similarity 96.5%; Pred. No. 9, 2e-145;
Matches 278; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 MSSNMWIDAFTEEEELALDAIEASYNFSSSSSSSAAPVQATTSVGHEDDNP 60
DB 1 MSSNMWIDAFTEEEELALDAIEASYNFSSSSSSSAAPVQATTSVGHEDDNP 60
QY 61 NTRQLPRSTSTSTSYKRPFLSRCARNPFPARFGRLYKSTATEVDKRAQLIKV 120
DB 61 NTRQLPRSTSTSTSYKRPFLSRCARNPFPARFGRLYKSTATEVDKRAQLIKV 120
QY 121 KRDESGIAFVGLDIEWRPSEFRKGLPGKVAIVQICVDNVCVMIHFHSGIPQSL 180
DB 121 KRDESGIAFVGLDIEWRPSEFRKGLPGKVAIVQICVDNVCVMIHFHSGIPQSL 180
QY 181 DSTLVKVGIGIDGVSUKLPHDYGVSITKVEDISDLANKTIGDGRKGLASTLET 240
DB 181 DSTLVKVGIGIDGVSUKLPHDYGVSITKVEDISDLANKTIGDGRKGLASTLET 240
QY 241 LKPNRIRLGNMEFFYPLSKOOLQYAATDAVASMHLKYLKDLDPVAGS 288
DB 241 LKPNRIRLGNMEFFYPLSKOOLQYAATDAVASMHLKYLKDLDPVAGS 288

RESULT 6
AAG52859
ID AAG52859 standard; Protein; 197 AA.
XX
AC AAG52859;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 67238.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
PD
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 9905-0130449.
PR 23-APR-1999; 9905-0130510.
PR 23-APR-1999; 9905-0130891.
PR 28-APR-1999; 9905-0131449.
PR 30-APR-1999; 9905-0132048.
PR 30-APR-1999; 9905-0132407.
PR 04-MAY-1999; 9905-0132484.
PR 05-MAY-1999; 9905-0132485.
PR 06-MAY-1999; 9905-0132486.
PR 07-MAY-1999; 9905-0132487.
PR 11-MAY-1999; 9905-0132863.
PR 14-MAY-1999; 9905-0134256.
PR 14-MAY-1999; 9905-0134218.
PR 14-MAY-1999; 9905-0134219.
PR 14-MAY-1999; 9905-0134221.
PR 14-MAY-1999; 9905-0134370.
PR 18-MAY-1999; 9905-0134768.
PR 19-MAY-1999; 9905-0134941.
PR 20-MAY-1999; 9905-0135124.
PR 21-MAY-1999; 9905-0135353.
PR 24-MAY-1999; 9905-0135629.
PR 25-MAY-1999; 9905-0136021.
PR 27-MAY-1999; 9905-0136392.
PR 28-MAY-1999; 9905-0136782.
PR 01-JUN-1999; 9905-0137222.
PR 03-JUN-1999; 9905-0137528.
PR 04-JUN-1999; 9905-0137502.
PR 07-JUN-1999; 9905-0137724.
PR 08-JUN-1999; 9905-0138094.
PR 10-JUN-1999; 9905-0138540.
PR 10-JUN-1999; 9905-0138847.
PR 14-JUN-1999; 9905-0139119.
PR 16-JUN-1999; 9905-0139452.
PR 16-JUN-1999; 9905-0139453.
PR 17-JUN-1999; 9905-0139492.
PR 18-JUN-1999; 9905-0139454.
PR 18-JUN-1999; 9905-0139455.
PR 18-JUN-1999; 9905-0139456.
PR 18-JUN-1999; 9905-0139457.
PR 18-JUN-1999; 9905-0139458.
PR 18-JUN-1999; 9905-0139459.
PR 18-JUN-1999; 9905-0139460.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139750.
PR 18-JUN-1999; 9905-0139763.
PR 21-JUN-1999; 9905-0139817.
PR 22-JUN-1999; 9905-0139899.
PR 23-JUN-1999; 9905-0140353.
PR 23-JUN-1999; 9905-0140354.
PR 24-JUN-1999; 9905-0140695.
PR 28-JUN-1999; 9905-0140823.
PR 29-JUN-1999; 9905-0140991.
PR 30-JUN-1999; 9905-0141287.
PR 01-JUL-1999; 9905-0141642.
PR 01-JUL-1999; 9905-0142154.
PR 02-JUL-1999; 9905-0142055.
PR 06-JUL-1999; 9905-0142390.
PR 08-JUL-1999; 9905-0142803.
PR 09-JUL-1999; 9905-0142920.
PR 12-JUL-1999; 9905-0142977.
PR 13-JUL-1999; 9905-0143542.
PR 14-JUL-1999; 9905-0143624.
PR 15-JUL-1999; 9905-0144005.
PR 16-JUL-1999; 9905-0144086.
PR 16-JUL-1999; 9905-0144085.
PR 19-JUL-1999; 9905-0144325.
PR 19-JUL-1999; 9905-0144331.
PR 19-JUL-1999; 9905-0144332.
PR 19-JUL-1999; 9905-0144333.
PR 19-JUL-1999; 9905-0144334.
PR 19-JUL-1999; 9905-0144335.

PR 20-JUL-1999; 9905-0144352.
PR 20-JUL-1999; 9905-0144632.
PR 20-JUL-1999; 9905-0144684.
PR 21-JUL-1999; 9905-0144814.
PR 21-JUL-1999; 9905-0145086.
PR 21-JUL-1999; 9905-0145088.
PR 22-JUL-1999; 9905-0145085.
PR 22-JUL-1999; 9905-0145087.
PR 22-JUL-1999; 9905-0145089.
PR 22-JUL-1999; 9905-0145192.
PR 23-JUL-1999; 9905-0145145.
PR 23-JUL-1999; 9905-0145218.
PR 23-JUL-1999; 9905-0145224.
PR 26-JUL-1999; 9905-0145276.
PR 27-JUL-1999; 9905-0145913.
PR 27-JUL-1999; 9905-0145918.
PR 27-JUL-1999; 9905-0145919.
PR 28-JUL-1999; 9905-0145951.
PR 02-AUG-1999; 9905-0146386.
PR 02-AUG-1999; 9905-0146388.
PR 02-AUG-1999; 9905-0146389.
PR 03-AUG-1999; 9905-0147038.
PR 04-AUG-1999; 9905-0147204.
PR 04-AUG-1999; 9905-0147302.
PR 05-AUG-1999; 9905-0147192.
PR 05-AUG-1999; 9905-0147260.
PR 06-AUG-1999; 9905-0147303.
PR 06-AUG-1999; 9905-0147416.
PR 09-AUG-1999; 9905-0147493.
PR 10-AUG-1999; 9905-0147935.
PR 10-AUG-1999; 9905-0148171.
PR 11-AUG-1999; 9905-0148319.
PR 12-AUG-1999; 9905-0148341.
PR 13-AUG-1999; 9905-0148565.
PR 16-AUG-1999; 9905-0148684.
PR 17-AUG-1999; 9905-0149368.
PR 18-AUG-1999; 9905-0149175.
PR 18-AUG-1999; 9905-0149426.
PR 20-AUG-1999; 9905-0149722.
PR 20-AUG-1999; 9905-0149723.
PR 20-AUG-1999; 9905-0149929.
PR 23-AUG-1999; 9905-0149902.
PR 23-AUG-1999; 9905-0149930.
PR 23-AUG-1999; 9905-0150566.
PR 25-AUG-1999; 9905-0150566.
PR 26-AUG-1999; 9905-0150884.
PR 27-AUG-1999; 9905-0151065.
PR 27-AUG-1999; 9905-0151066.
PR 27-AUG-1999; 9905-0151080.
PR 30-AUG-1999; 9905-0151303.
PR 31-AUG-1999; 9905-0151438.
PR 01-SEP-1999; 9905-0151930.
PR 07-SEP-1999; 9905-0152363.
PR 10-SEP-1999; 9905-0153070.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154018.
PR 16-SEP-1999; 9905-0154039.
PR 20-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0155486.
PR 24-SEP-1999; 9905-0155659.
PR 28-SEP-1999; 9905-0156458.
PR 29-SEP-1999; 9905-0156596.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.
PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158232.
PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159294.
PR 14-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159330.

PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 68.9%; Score 1028; DB 21; Length 197;
Best Local Similarity 100.0%; Pred. No. 1.3e-101;
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 92 MRFGRILYKSTATEVDRKAMQLIKVLDPTKRDESGIAFVGJDIEMRPSFRKGVLPKQVAT 151
DB 1 MRFGRILYKSTATEVDRKAMQLIKVLDPTKRDESGIAFVGJDIEMRPSFRKGVLPKQVAT 60
QY 152 VOICVDNVCYCVMIHFHSGITQSOLOHLEIDSTLYKVGIGIGDSYKLFHDYGVSTKQVED 211
DB 61 VOICVDNVCYCVMIHFHSGITQSOLOHLEIDSTLYKVGIGIGDSYKLFHDYGVSTKQVED 120
QY 212 LSDLANOKIGDKMKGLASTETLYCKELKPNRIKGNWFEYPLSKOOLQYATDAYAS 271
DB 121 LSDLANOKIGDKMKGLASTETLYCKELKPNRIKGNWFEYPLSKOOLQYATDAYAS 180
QY 272 WHLYKVLKLDLPDAVSGS 288
DB 181 WHLYKVLKLDLPDAVSGS 197

RESULT 7

AAG24978

ID AAG24978 standard; Protein; 197 AA.
XX AAG24978;
AC AAG24978;
XX AAG24978;
DT 17-OCT-2000 (first entry)
XX AAG24978;
DE Arabidopsis thaliana protein fragment SEQ ID NO: 28857.
XX AAG24978;
KW Protein identification: signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
XX Arabidopsis thaliana.
PN EPI033405-A2.
XX EPI033405-A2.
PD 06-SEP-2000.
XX 06-SEP-2000.
PF 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-2000; 2000EP-0301439.
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.

PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144353.
PR 20-JUL-1999; 99US-0144354.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145313.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148365.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.

PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161820.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 66.8%; Score 996; DB 21; Length 197;
Best Local Similarity 97.0%; Pred. No. 3, 5e-98;

Matches 191; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 92 MRFGRILYSKATETEDKRAMOLIKVLDTRKDSGTAFAVGLDIEMPSFRKGLPKGVAT 151
Db 1 MRFGRILYSKATETEDKRAMOLIKVLDTRKDSGTAFAVGLDIEMPSFRKGLPKGVAT 60
QY 152 VOICVDNVCYDVNHFHSGIPQSLQHLIEDSTLVKVGIGIDGSVKLFHDYGVSIKDED 211
Db 61 VOICVDNVCYDVNHFHSGIPQSLQHLIEDSTLVKVGIGIDGSVKLFHDYGVSIKDED 120
QY 212 LSDLANOKIGGDKKKNLSTETLVCKELKPRIRILGNWEPYPLSKOOLQYATATAYAS 271
Db 121 LSDLANOKIGGDKKKNLSTETLVCKELKPRIRILGNWEPYPLSKOOLQYATATAYAS 180
QY 272 WHLYKVKLDLPDAVSGS 288
Db 181 WHLYKVKLDLPDAVSDS 197

RESULT 8
AAG24979
ID AAG24979 standard; Protein; 177 AA.
XX
AC AAG24979;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 28858.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
PN EP1033405-A2.
XX
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0125348.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139453.
PR 16-JUN-1999; 99US-0139455.
PR 17-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.

```

PR 29-DEC-1995; 95US--0009409.
PR 29-DEC-1995; 95US--0580539.
PR 30-JAN-1996; 96US--0010835.
PR 30-JAN-1996; 96US--0594242.
XX
PA (DARW-) DARWIN MOLECULAR CORP.
PA (OSHI/) OSHIMA J.
XX
PI Fu Y, Mulligan J, Oshima J, Schellenberg GD, Yu C;
XX
XX MPI: 1997-363671/33.
DR N-PSDB: AAX83001.
DR
XX
XX Isolated nucleic acid molecule encoding the WRN gene product
PT useful for detection and treatment of Werner's syndrome, and related
PT diseases
XX
PS Claim 10; Fig 2B; 153pp; English.
XX
XX This sequence represents the human WRN gene product which is associated
CC with Werner's syndrome. The products can be used for the detection and
CC treatment of Werner's syndrome (WS), an autosomal recessive disorder
CC with a complex phenotype, as well as related diseases.
XX
XX Sequence 1432 AA;
XX

```

Query Match	23.68;	Score 352;	DB 18;	Length 1432;
Best Local Similarly	37.78;	Pred. No. 7.6e-28;		
Matches 75; Conservative	41;	Mismatches 75;	Indels 8;	Gaps 3;

QY 88 NFPMFEGGIIYSKATVDRRAQOLIKVDPTRKDEEGIAFVGLDIEDMRSPFKGVLPG 147
 Db 43 DLFELFETGSIYVSTADSCPSLSE-----DISMSLDGDDVGEFDMEMPLYNRGL-G 95
 QY 148 KVATVQICVDSNCTCMHIIHFSGI-POSQIHLIEDSTLYKVGIGIIDDQVKLFHFIDYVI 206
 Db 96 KVALIDICVSESCTCYFHVSSMVFPPQGIKMLEAKKAGVGEIGGOMKLLDHPDIFKL 155
 QY 207 KDVEDLSIDLANOKIIGGDKRWGLASTETFLVCKELKLPRIIRLGMWEFYPLSKOOLQYAA 266
 Db 156 KNFVELITDVAANKLCTETFSVNSLYVKHLGLQOLKDKSICSMMSKSPFLTEDOKLVAAT 215

XX	RESULT 10
AAAM50935	
ID	AAAM50935 standard; Protein; 1432 AA.
XX	
AC	AAAM50935;
XX	
DT	15-MAY-2002 (first entry)
XX	
DE	Human Werner syndrome protein (3'-5' exonuclease domain).
XX	
KM	Enzyme; 3'-5' exonuclease; human; gene expression;
KW	Werner syndrome protein; gene silencing; transgenic plant.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Domain
FT	Peptide
FT	Peptide
FT	Peptide
FT	Peptide
FT	Peptide
FT	Peptide
XX	
NN	WO200210362-A2.

```

XX 07-FEB-2002.
PD
XX
XX 30-JUL-2001; 2001WO-EP08825.
PF
XX
XX 01-AUG-2000; 2000US-222202P.
PR
XX
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.
PA (NOVS ) NOVARTIS FORSCHUNGSSTIFTUNG ZWEIGNIEDERL.
XX
XX Levin JZ, Phillips KL, Budziszewski GJ, Meins F, Glazov EA;
XX WPI; 2002-217119/27.
DR N-PSDB: ABA91803.
XX
XX Regulating gene expression in plants for controlling gene silencing,
PT comprises altering the transcription or translation of an endonuclease
PT nucleotide sequence encoding a polypeptide comprising an exonuclease
PT domain
XX
XX Claim 14; Page 86-91; 102pp; English.
PS
XX The present is that of human Werner syndrome protein (GenePept
CC accession AAF06162), which includes an N-terminal 3'-5' exonuclease
CC domain. The protein was identified in a homology screening with an
CC Arabidopsis thaliana 3'-5' exonuclease domain (see AAM50927). The
CC 3'-5' exonuclease domains consists of 3 sequence motifs, termed
CC Exoi, Exoii and Exoiii. Nucleotide (see ABA91795-807) and
CC polypeptide (see AAM50927-39) sequences for 3'-5' exonuclease
CC domains were identified using different screening methods and search
CC algorithms. The invention encompasses the suppression or increase
CC of gene silencing in plants. This is achieved by altering the
CC expression in the plant cell of a nucleotide sequence encoding a
CC polypeptide, such as the present sequence, having an exonuclease
CC domain, preferably a 3'-5' exonuclease domain, and especially
CC an RNase D related domain. Expression of the nucleotide sequence
CC is altered by altering its transcription or translation. It is
CC reduced e.g. by sense suppression, antisense suppression, homologous
CC recombination, the use of ribozymes, dominant-negative mutants,
CC aptamers, zinc finger proteins, double-stranded RNA, or insertional,
CC point or deletion mutagenesis, and increased by overexpression.
CC The methods are useful for altering or stabilizing the expression of
CC a nucleotide sequence of interest in a plant cell or plant.
CC Suppressing or decreasing expression of the nucleic acid molecule
CC results in decreased levels of post transcriptional gene silencing
CC and improved expression of genes of interest. The 3'-5'
CC exonuclease domain polypeptides are also useful for identifying
CC compounds capable of altering 3'-5' exonuclease activity.
XX
XX Sequence 1432 AA;
SQ
Query Match 23.6%; Score 352; DB 23; Length 1432;
Best Local Similarity 37.7%; Pred. No. 7.6e-28;
Matches 75; Conservative 41; Mismatches 75; Indels 8; Gaps 3;
QY 88 NFPMARFGGRILYKSTATEVVKRAMQLIKVDTRKDESGIAFVGIDIEMRPSFRKGYLP 147
DB 43 DLPLEFTGSLIVYSDSCFLSE-----DISMSISDGVYGFDMEMPLYNKGL-G 95
QY 148 KVAATVQICVDSNYCDVMHIFHSGI-POSLQHLIEDSTLVKVGIGIDGSKVLFFHDYGVSI 206
DB 96 KVALIQCVSESKCYLRFHVSSMSYFPOGLKMLENKAVKKAGVIEGQWMLLNDPDKL 155
QY 207 KDVEDLSDLANQKIGGDKKMGASLTETLVCKELKPRIRLGNWERYPLSKQOLQYAT 266
DB 156 KNFVELTDVANKKLCETWLSNLSVKHLGKOLLKDKSIRCSNWSKPLETDOKLYAAT 215
QY 267 DAYASWHLXYKVLKDLDPAY 285
DB 216 DAYAGFIITRYKLEITLDTIV 234

```

RESULT 11

```

AAM59454
ID AAM59454 standard; Protein; 1401 AA.
XX
AC AAM59454;
XX
DT 28-AUG-1998 (first entry)
XX
DE Mouse WRN helicase protein.
XX
KM Werner's syndrome; WRN; helicase; murine; diagnosis.
XX
OS Mus sp.
XX PN JP10146188-A.
XX
PD 02-JUN-1998.
XX
PF 15-NOV-1996; 96JP-0304721.
XX
PR 15-NOV-1996; 96JP-0304721.
XX
PA (EIJT-) EIJIN KENKYUSHO KK.
XX
DR WPI; 1998-369863/32.
DR N-PSDB: AAV35114, AAV35115.
XX
XX New mouse gene encoding protein - used for diagnosis of human
PT Werner's syndrome
PT
PS Claim 1; Page 18-22; 30pp; Japanese.
XX
XX This sequence represents a novel mouse WRN helicase protein which is
CC associated with Werner's syndrome. The mouse WRN gene is useful for
CC the diagnosis of human Werner's syndrome.
XX
XX Sequence 1401 AA;
SQ
Query Match 22.7%; Score 339; DB 19; Length 1401;
Best Local Similarity 36.6%; Pred. No. 1.8e-26;
Matches 74; Conservative 42; Mismatches 72; Indels 14; Gaps 5;
QY 88 NFPMARFGGRILYKSTATEVVKRAMQLIKVDTRKDESGIAFVGIDIEMRPSFRKGYLP 147
DB 37 NLPLEFEGGSIVSYEASD-----CSFLSEDISMLSDGV--VGFDMEMPLIYK----PG 86
QY 148 K---VAIVQICVDSNYCDVMHIFHSGI-POSLQHLIEDSTLVKVGIGIDGSKVLFFHDY 203
DB 87 KRSRYAVIOLCVSEKCYLRFHISMSYFPOGLKMLENKSIKAGVIEGQWMLLNDPDKL 146
QY 204 VSIRDVEDLSDLANQKIGGDKKMGASLTETLVCKELKPRIRLGNWERYPLSKQOLQY 263
DB 147 VKLESFVELTDVANKKLCETWLSNLSVKHLGKOLLKDKSIRCSNWSKPLETDOKLY 206
QY 264 AATDAYASWHLXYKVLKDLDPAY 285
DB 207 AATDAYAGFIITRYKLEITLDTIV 228

```

RESULT 12

```

AAM97841
ID AAM97841 standard; Protein; 1401 AA.
XX
AC AAM97841;
XX
DT 07-JUN-1999 (first entry)
XX
DE Murine mWRN polypeptide (WRN homologue).
XX
KM mWRN gene; WRN gene; homologue; Werner's syndrome; mouse;
KM life span; aging.
XX
OS Mus sp.
XX

```

PN W09906543-A1.
 XX 11-FEB-1999.
 PD
 XX
 PF 03-AUG-1998: 98MO-US16081.
 XX
 PR 04-AUG-1997: 97US-0054629.
 XX
 PA (MASI) MASSACHUSETTS INST TECHNOLOGY.
 XX
 PI Guarante LP, Lombard DB, Sinclair DA.
 XX
 DR WPI, 1999-153773/13.
 DR N-PSDB; AAX24302.
 XX
 PT Identifying agents that inhibit replication and accumulation of DNA
 PT circles - as potential agents for increasing the life span of
 PT animals
 XX
 PS Claim 13: Page 69-72; 78pp; English.
 XX
 CC This is the amino acid sequence of a polypeptide encoded by mWRN
 CC see AAX24302), the murine homologue of the human WRN gene associated
 CC with Werner's syndrome. A claimed assay for identifying a compound
 CC which extends life span comprises administering a compound to be
 CC tested to a mammal with a suppressed level of mWRN, and identifying
 CC slowing of at least one of the normal phenotypes of ageing in the
 CC mammal. A compound identified by this method inhibits the
 CC replication and/or accumulation of rDNA circles in a cell. The
 CC assay can be performed using a mWRN knockout mouse (also claimed).
 CC mWRN nucleic acids can be used to produce mWRN polypeptides and as
 CC diagnostic probes and primers. mWRN polypeptides can be used to
 CC raise antibodies (for inhibiting activity of mWRN or for determining
 CC it in tissues). The assay is based on the observation that in yeast
 CC cells, accumulation of ribosomal circular DNA is responsible for
 CC age-related enlargement and fragmentation of the nucleolus. A
 CC mutation in the yeast WRN homologue Sgs1 causes premature ageing,
 CC suggesting a common ageing mechanism in all eukaryotes. Insight
 CC into the ageing process in model systems can provide insight into
 CC ageing in humans.
 CC
 SQ Sequence 1401 AA;
 Query Match 22.7%; Score 339; DB 20; Length 1401;
 Best Local Similarity 36.6%; Pred. No. 1.8e-26;
 Matches 74; Conservative 42; Mismatches 72; Indels 14; Gaps 5;
 QY 88 NPPAMRFGGRILYSKATATEVDKRAMQLIVLDTKRDESGIAFYGLDIEMRPSRKGVLP 147
 DB 37 NLPLEFPGSIYSYASD----CSFLSEDISMRSDGV--VGFDMEWPPIYK----PG 86
 QY 148 K--VATVOICVDSNYCDVMHIFHSGI-POSLOHLIEDSTLVKVGIGIGDSVKLFPHDYG 203
 DB 87 KRSRAVAVIOLCVSESKCYLFHISMSVFPQGLKMLENKSIRKAGVIGEDQKLLRPD 146
 QY 204 VSIKDVEDSDLANOKIGDKKGLASTETLVCKELKPNRIKLNMEFYPLSKOOLQY 263
 DB 147 VKLESEVELTDVANEKLCATETWSLNGLYKHYGLKQLLDKDSIRCSNMSNFPLEDOKILY 206
 QY 264 AATDAYASWHLKYVKLDPDAV 285
 DB 207 AATDAYAGLIIYOKLGNLGDV 228
 RESULT 13
 ID AAY14519 standard; Protein; 1401 AA.
 AC AAY14519;
 XX
 DT 31-AUG-1999 (first entry)
 XX
 DE Mouse WRN gene product.

XX
 KW Human; WRN; Werner's syndrome; detection; diagnosis; autosomal;
 KM recessive disorder; phenotype.
 XX
 OS Mus musculus.
 XX
 PN W09724435-A1.
 XX
 PD 10-JUL-1997.
 XX
 PF 30-DEC-1996: 96MO-US20785.
 XX
 PR 12-APR-1996: 960S-0632175.
 PR 29-DEC-1995: 95US-0009409.
 PR 29-DEC-1995: 95US-0580539.
 PR 30-JAN-1996: 96US-0010835.
 PR 30-JAN-1996: 96US-0594242.
 XX
 PA (DARN-) DARWIN MOLECULAR CORP.
 PA (OSHI/) OSHIMA J.
 XX
 PI Fu Y, Mulligan J, Oshima J, Schellenberg GD, Yu C;
 XX
 DR WPI, 1997-363671/33.
 DR N-PSDB; AAX83004.
 XX
 PT Isolated nucleic acid molecule encoding the WRN gene product
 PT useful for detection and treatment of Werner's syndrome, and related
 PT diseases
 XX
 PS Claim 10: Fig 6; 153pp; English.
 XX
 CC This sequence represents the mouse WRN gene product. The corresponding
 CC human gene product (AAY14517) is associated with Werner's syndrome. The
 CC products can be used for the detection and treatment of Werner's
 CC syndrome (WS), an autosomal recessive disorder with a complex phenotype,
 CC as well as related diseases.
 CC
 SQ Sequence 1401 AA;
 Query Match 22.1%; Score 330; DB 18; Length 1401;
 Best Local Similarity 36.0%; Pred. No. 1.7e-25;
 Matches 72; Conservative 43; Mismatches 71; Indels 14; Gaps 5;
 QY 88 NPPAMRFGGRILYSKATATEVDKRAMQLIVLDTKRDESGIAFYGLDIEMRPSRKGVLP 147
 DB 37 NLPLEFPGSIYSYASD----CSFLSEDISMRSDGV--VGFDMEWPPIYK----PG 86
 QY 148 K--VATVOICVDSNYCDVMHIFHSGI-POSLOHLIEDSTLVKVGIGIGDSVKLFPHDYG 203
 DB 87 KRSRAVAVIOLCVSESKCYLFHISMSVFPQGLKMLENKSIRKAGVIGEDQKLLRPD 146
 QY 204 VSIKDVEDSDLANOKIGDKKGLASTETLVCKELKPNRIKLNMEFYPLSKOOLQY 263
 DB 147 VKLESEVELTDVANEKLCATETWSLNGLYKHYGLKQLLDKDSIRCSNMSNFPLEDOKILY 206
 QY 264 AATDAYASWHLKYVKLDPD 283
 DB 207 AATDAYAGLIIYOKLGNLGD 226
 RESULT 14
 ID ABB71575 standard; Protein; 346 AA.
 AC ABB71575;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide seq ID NO 41517.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 10:21:45 ; Search time 19 Seconds
(without alignments)
1457.195 Million cell updates/sec

Title: US-09-896-186B-24

Perfect score: 1491

Sequence: 1 MSSSNWIDDAFTFEELALD.....YASWHLKYVLKDLDPDAVSGS 288

Scoring table:

BIOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	1439.5	96.5	313	2	T05256 hypothetical prote
2	339.5	22.8	1436	2	T14895 DNA helicase 1 - A
3	339	22.7	1401	2	T30247 Werner syndrome pr
4	330	22.1	1401	2	T17452 Werner syndrome pr
5	179.5	12.0	582	2	H96604 probable 3'-5' exo
6	158	10.6	123	2	E84752 hypothetical prote
7	136.5	9.2	910	2	S40930 hypothetical prote
8	127.5	8.6	239	2	A84777 hypothetical prote
9	112.5	7.5	217	2	T02548 hypothetical prote
10	101	6.8	816	2	T19246 hypothetical prote
11	97.5	6.5	445	2	E72674 hypothetical prote
12	96.5	6.5	488	2	F97039 hypothetical prote
13	95	6.4	508	1	A36121 aromatase (EC 1.14
14	95	6.4	4589	2	T14914 dynein beta heavy
15	94.5	6.3	416	2	C71620 protein with Egl-1
16	93.5	6.2	1525	1	GNWVS5 genome polyprotein
17	93	6.2	2224	1	KFNHUS coagulation factor
18	93	6.2	3079	1	RGBY12 probable GTPase-ac
19	92	6.2	237	2	H84732 hypothetical prote
20	92	6.2	758	2	T39210 3-isopropylmalate
21	91.5	6.1	473	2	S36553 L2 protein - human
22	91.5	6.1	781	2	T41551 hypothetical prote
23	91	6.1	503	2	S13912 aromatase (EC 1.14
24	90	6.0	472	2	T47436 protein kinase-11k
25	89.5	6.0	930	2	E64098 DNA-directed DNA p
26	89.5	6.0	1324	2	T01508 mismatch repair en
27	88	5.9	269	2	I50504 nk2.2 protein - ze
28	88	5.9	317	2	A97511 hypothetical prote
29	88	5.9	317	2	AH2729 agmatinase [import

30	88	5.9	417	2	E82095 gamma-glutamyl pho
31	88	5.9	709	2	S40926 hypothetical prote
32	88	5.9	784	2	C88558 protein ZK1098.3 l
33	87.5	5.9	334	2	T03556 probable D-ribose-
34	87.5	5.9	422	2	E96753 hypothetical prote
35	87.5	5.9	547	2	T40342 signal recognition
36	87.5	5.9	589	2	E88492 protein T07E3.1 [1
37	87	5.8	178	2	E72450 hypothetical prote
38	87	5.8	392	2	T44362 acetyl-CoA C-acety
39	87	5.8	579	2	I40371 methyltransferase
40	87	5.8	3137	1	A37797 collagen alpha 3(V
41	86.5	5.8	346	1	BRECPR phosphate-repressi
42	86.5	5.8	360	2	T00882 hypothetical prote
43	86.5	5.8	470	2	T10193 glycerol-3-phospha
44	86.5	5.8	496	2	S61966 hypothetical prote
45	86.5	5.8	593	2	B89844

ALIGNMENTS

RESULT 1

T05256 hypothetical protein F18A5.260 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999

C:Accession: T05256
R:Bevan, M.; Weber, N.; Grueninger, D.; Schmidheini, T.; Bancroft, I.; Mewes, H.W.; M
submitted to the Protein Sequence Database, February 1999

A:Reference number: 215405

A:Accession: T05256

A:Molecule type: DNA

A:Residues: 1-313 <BEV>

A:Cross-References: EMBL:AL035528

A:Experimental source: cultivar Columbia; BAC clone F18A5

C:Genetics:

A:Map position: 4

A:Introns: 86/1; 143/1; 186/3; 239/3

A>Note: F18A5.260

Query Match 96.5%; Score 1439.5; DB 2; Length 313;
Best Local Similarity 95.3%; Pred. No. 5.6e-117;
Matches 281; Conservative 1; Mismatches 4; Indels 9; Gaps 1;

QY	1	MSSSNWIDDAFTFEELALDAIDAEASYNFSRSSSSSSAAPTVOATTSVGHEDPQIPN	60
DB	1	MSSSNWIDDAFTFEELALDAIDAEASYNFSRSSSSSSAAPTVOATTSVGHEDPQIPN	60
QY	61	NIRROLPRSTSTSTSKRPPLSRCRARNPAMRFGGRILYKTATEVDKRAMQLIKVLDT	120
DB	61	NIRROLPRSTSTSTSKRPPLSRCRARNPAMRFGGRILYKTATEVDKRAMQLIKVLDT	120
QY	121	KRDESGIAFVGDIEMRPSFRKGVLPKQATVOICVDSNYCDVMHIFHSIGIPQSLQHLT	180
DB	121	KRDESGIAFVGDIEMRPSFRKGVLPKQATVOICVDSNYCDVMHIFHSIGIPQSLQHLT	180
QY	181	DSTLYKVGIGIDGYSVKLFHDYGVSIKQVEDLSIDLANOKTGDGKKWGLASTRETYCKEL	240
DB	181	DSTLYKVGIGIDGYSVKLFHDYGVSIKQVEDLSIDLANOKTGDGKKWGLASTRETYCKEL	240
QY	241	LKPNRIKGNMFEYPLSKOOLYATADVASWHLKYV-----LKDLPAVS	286
DB	241	LKPNRIKGNMFEYPLSKOOLYATADVASWHLKYVTTTKNHLTLINDLEAKIS	295

RESULT 2

T14895 DNA helicase 1 - African clawed frog

C:Species: Xenopus laevis (African clawed frog)

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: T14895

R:Yan, H.; Chen, C.Y.; Kobayashi, R.; Newport, J.

Nature Genet. 19, 375-378, 1998

A>Title: Replication focus-forming activity 1 and the Werner syndrome gene product.
A:Reference number: Z18255; MID:98361165; PMID:9697700
A:Accession: T14895
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 11436 <XAN>
A:Cross-references: EMBL:AF067418; NID:g3420290; PID:g3420291; PIDN:AAC63512.1
C:Genetics:
A:Gene: FRA-1

Query Match 22.8%; Score 339.5; DB 2; Length 1436;
Best Local Similarity 35.0%; Pred. No. 8.5e-21;

Matches 82; Conservative 43; Mismatches 84; Indels 25; Gaps 7;

QY 61 NRRQPLRSITSSTYSKRF---PLSRCAR---NFPAMRGRILYSKTATEV-----D 108
DB 3 SLQRKLPKMSVQOQDRIDDAKSKCKNILEDNLPKMFNSIYYSINCSLSED 62
QY 109 KRAMOLIKVLDTRKDESGIAFVGLDIEMRPSFRKGYLPCKVATVOICVDSNYDVMTF- 167
DB 63 IRSSSL-----BEDV--LGFDIEMPVYTKG-KTGKVALIQVSEKCKYLFHISP 110
QY 168 HSGIPQSLQHLIEDSLVAVGIGIDDSVKLFHDYGVSTVDVEDSLANOKIGDKKMG 227
DB 111 MAFEPGLKRLKLEDSYKRVGVEGQWKLMSDYELKLTGTELSMANOKRCKEKT 170
QY 228 LSLATFLVCKELKFNRIKGMWEFPLSKOOLQYATDAYASWMLYKLDL 281
DB 171 ENGLIHLEFQELYKRSYRCSNMDIFLLEDDQKLVATDAYAGLITKLEEM 224

RESULT 3
T30247
Werner syndrome protein type1 - mouse

N:Alternate names: Wtn type1 protein
C:Species: Mus musculus (house mouse)

C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 21-Jul-2000
C:Accession: T30247
R:Imamura, O.; Ichikawa, K.; Yamabe, Y.; Goto, M.; Sugawara, M.; Furutachi, Y.

A>Title: Cloning of a mouse homologue of the human Werner syndrome gene and assignment
A:Reference number: Z20785; MID:97288537; PMID:9143515
A:Accession: T30247

A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1401 <TMA>

A:Cross-references: EMBL:DB6526; NID:g2130972; PIDN:BA420269.1; PID:g2130973
A:Experimental source: strain BALB/c; testis/spleen
C:Genetics:

A:Gene: WRN type1
A:Map position: 8A4

Query Match 22.7%; Score 339; DB 2; Length 1401;
Best Local Similarity 36.6%; Pred. No. 9.1e-21;

Matches 74; Conservative 42; Mismatches 72; Indels 14; Gaps 5;

QY 88 NFPAMRGRILYSKTATEVDKRAMOLIKVLDTRKDESGIAFVGLDIEMRPSFRKGYLP 147
DB 37 NLPFLFPGSIYSYASD---CSFLSEDISMRISDGV--VGFMEMPPIYK---PG 86
QY 148 K---VATVOICVDSNCDVMTFHSGI-POSTLOHLEDSTLVAVGIGIDDSVKLFHDY 203
DB 87 KRSRAVAVQLCVSESKCYLHISMSVFPQGLKMLLENKSIKAGVIGEDDOKRLRDFD 146
QY 204 VSIRKVEDSLANOKIGDKKGLASLFTETLVCKELKFNRIKGMWEFPLSKOOLQY 263
DB 147 VKLESFVELTDVANEKLCATETWSLNGLVKLVGKQLDKSIRCSNWSNFPLEDDOKLY 206
QY 264 AATDAYASWMLYKVLKLDLPD 285
DB 207 AATDAYAGLITVYKLGNDTV 228

RESULT 4
T17452
Werner syndrome protein - mouse

N:Alternate names: Wtn protein
C:Species: Mus musculus (house mouse)

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T17452
R:Pepper, B.W.; Gayle, M.; Brady, W.; Swartz, A.; Gillett, L.A.; Altsch, R.S.; Mullig

A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MID:21016719; PMID:11130712
A:Accession: H96604

A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1582 <STO>

A:Cross-references: GB:AE005173; NID:g11094727; PIDN:AAG29662.1; GSPDB:GN00141
C:Genetics:
A:Gene: F14G9.8
A:Map position: 1

Query Match 22.1%; Score 330; DB 2; Length 1401;
Best Local Similarity 36.0%; Pred. No. 5.5e-20;

Matches 72; Conservative 43; Mismatches 71; Indels 14; Gaps 5;

QY 88 NFPAMRGRILYSKTATEVDKRAMOLIKVLDTRKDESGIAFVGLDIEMRPSFRKGYLP 147
DB 37 NLPFLFPGSIYSYASD---CSFLSEDISMRISDGV--VGFMEMPPIYK---PG 86
QY 148 K---VATVOICVDSNCDVMTFHSGI-POSTLOHLEDSTLVAVGIGIDDSVKLFHDY 203
DB 87 KRSRAVAVQLCVSESKCYLHISMSVFPQGLKMLLENKSIKAGVIGEDDOKRLRDFD 146
QY 204 VSIRKVEDSLANOKIGDKKGLASLFTETLVCKELKFNRIKGMWEFPLSKOOLQY 263
DB 147 VKLESFVELTDVANEKLCATETWSLNGLVKLVGKQLDKSIRCSNWSNFPLEDDOKLY 206
QY 264 AATDAYASWMLYKVLKLDLP 283
DB 207 AATDAYAGLITVYKLGNDTV 226

RESULT 5
H96604

Probable 3'-5' exonuclease [Imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: H96604
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon

Chen, C.W.; Chung, M.K.; Cono, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Hutzat, U.
Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marzia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MID:21016719; PMID:11130712
A:Accession: H96604

A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1582 <STO>

A:Cross-references: GB:AE005173; NID:g11094727; PIDN:AAG29662.1; GSPDB:GN00141
C:Genetics:
A:Gene: F14G9.8
A:Map position: 1

Query Match 12.0%; Score 179.5; DB 2; Length 582;
Best Local Similarity 28.2%; Pred. No. 1.8e-07;

Matches 53; Conservative 36; Mismatches 66; Indels 33; Gaps 7;

QY 126 GIAFVGIDIEWRPSFRKGYLPKAVATVOICVDSN--YCDVMTFHSGI--IPQSLQHLIE 180
DB 377 GCRVVGIDCEWKFNRYIKGSKONKVSIMQISDTKPIFLDLIKLVNASEILNCLSHIIO 436

QY 181 DSTLVK-----GIGIDGSKLFHDYGVSIKDE-----LSDLANOKI 220
 Db 437 SKSTLKLVSLETDYDPDHKLSSGYNQCDIKQALSYG-DLKFERYDMLDQNVNPEF 495
 QY 221 GGDKKWGLASLETYLCKELLKPNIRLGNMEFFPLSKOOLQYATADYASWHLKVLKD 280
 Db 496 G-----GAGLTKKILIGLUNKTR--RNSDMEQRPLSQNQLLEYALDAVLIHRIHRVD 548
 QY 281 LPDAVSGS 288
 Db 549 HPDHSSS 556

RESULT 6

E84752

hypothetical protein At2g34110 (imported) - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: E84752

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: E84752

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-123 <STO>

A:Cross-references: GB:AE002093; NID:g2342731; PIDN:AA867629.1; GSPDB:GN00139

C:Genetics:

A:Molecule type: DNA

A:Gene: At2g34110

A:Map position: 2

Query Match 10.6%; Score 158; DB 2; Length 123;
 Best Local Similarity 79.5%; Pred. No. 1.6e-06;
 Matches 31; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 190 GIDGDSVKLFHDYGVSIKDEDSLANKIGGDKKML 228
 Db 82 GIDGDSVKLFHDYGVSIKDEDSLANKIGGDKKML 120

RESULT 7

S40930

hypothetical protein ZK1098.8 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Sep-1997

C:Accession: S40930

R:Thomas, K.

submitted to the EMBL Data Library, February 1992

A:Reference number: S40923

A:Accession: S40930

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-910 <THO>

A:Cross-references: EMBL:Z22176; NID:g297978; PID:g297986

C:Genetics:

A:Introns: 64/1; 336/2; 382/2; 447/2; 681/2; 810/1; 852/2

Query Match 9.2%; Score 136.5; DB 2; Length 910;
 Best Local Similarity 24.8%; Pred. No. 0.0017;
 Matches 77; Conservative 34; Mismatches 110; Indels 89; Gaps 13;

QY 12 TEEELALDAIEASYNFSRSSSSSAAPVQATTSVGHEDPQIENIRQLPREIT 71
 Db 333 TEKQ---IDAICPAIYLGIASSSS-----YQLP-NVWRDFRQPD 368
 QY 72 SSTYKRPPLSRCARANFPAMRFGGRILY-----SKATVEYDKRAMOLIKYID 119
 Db 369 SKLREAKELLVRKRLQV---LNAGEOLFVEENERRTOIHMKTESEMKN---YLCSEIK 421

QY 120 TKRDESGIAFVGLDIEMRPSFRKGVLPCKVATVQI-----CVD--SNYCDVMHI 166
 Db 422 SLSDERAPRYVGFDESEKRSNLTAVHDSXIALIQLFPRKCVWLVDCVELEKAMND----- 477
 QY 167 FHSGLPQSLQHLIEDSTLVKVGIGIDG-----SVKLFHDYGVSIKDEDSL 214
 Db 478 --DMWQKFAASRLFGSPVYGVFDMRNDADAMATIPALKSSMKI-----EDTKNAFIDLK 530
 QY 215 LANOKIGD-----KKWGLASLETYLCKELLKPNIRLGNMEFFPLSKOOLQYAT 266
 Db 531 LAENVCDDIMEILEDPRKTFKLADLTLYLGLDLKTE--QCSNQCRPLARKQIYVAL 588
 QY 267 DAYASWHLK 276
 Db 589 DAVVYVETFK 598

RESULT 8

A84777

hypothetical protein At2g36110 (imported) - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: A84777

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: A84777

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-239 <STO>

A:Cross-references: GB:AE002093; NID:g4678223; PIDN:AAD26968.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g36110

A:Map position: 2

Query Match 8.6%; Score 127.5; DB 2; Length 239;
 Best Local Similarity 25.1%; Pred. No. 0.0017;
 Matches 55; Conservative 33; Mismatches 90; Indels 41; Gaps 10;

QY 62 IRRQLPRSTSTSKRFPPLSRCARANFPAMRFGGRILYSKT-ATEVDKRAMQLIKYIDT 120
 Db 4 IRRQLPRSTSTSKRFPPLSRCARANFPAMRFGGRILYSKT-ATEVDKRAMQLIKYIDT 120
 QY 121 KRDESGIAFVGLDIEMRPSFRKGVLPCKVATVQIENIRQLPREIT 179
 Db 50 LRSSHPLV-VGLDVQWTP---GSDPDPDILQLCV-GNRCLILQLSHCKRIPEVLRSL 103
 QY 180 EDSTLVKVGIGIGIDGSKL---FHDYGV-SIKDEDSLANKIGGDKKWLAS 230
 Db 104 EDEITTFYGVWMSQOQKLERPHOLELWRLLDIRHYLPTRLNLSFEKI----- 153
 QY 231 LTELCKELLKPNIRLGNMEFFPLSKOOLQYATADY 269
 Db 154 VEECLGYKGVKDKCKEICHMSNMGARSLSHDQIVQASDVI 192

RESULT 9

T02548

hypothetical protein At2g32490 (imported) - Arabidopsis thaliana

N:Alternate names: hypothetical protein T26B15.5

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 16-Feb-2001

C:Accession: T02548; G84733

R:Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Syke

submitted to the EMBL Data Library, July 1998

A:Description: Arabidopsis thaliana chromosome II BAC T26B15 genomic sequence.

A:Reference number: Z14678

A:Accession: T02548

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-217 <R0U>
A:Cross-references: EMBL:AC004681; NID:g3298532; PID:g3298537
A:Experimental source: cultivar Columbia
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanhaken, S.E.; Umayam, L.; Tallon, L.
Eus, D.; Nierman, M.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: G84733
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-217 <STO>
A:Cross-references: GB:AE002093; NID:g3298537; PIDN:AAC25931.1; GSPDB:GN00139
C:Genetics:
A:Gene: T26B15.5; ATG32490
A:Map position: 2
A:Introns: 25/1

Query Match 7.5%; Score 112.5; DB 2; Length 217;
Best Local Similarity 25.4%; Pred. No. 0.03;
Matches 47; Conservative 38; Mismatches 77; Indels 23; Gaps 10;

QY 94 FCGRIYKATF-EVDKRAMQILKVDTRKDESGIAFVGLDIEWRSPFRKGVLPKVAAY 152
DB 31 FCGRIIVTVHTPSVIRRMHISIRFVSRRLSHPLV-VGLGVQWTP--RGSDDPDI--L 85
QY 153 QICVDNSYCDVMHIFH-SGIPSOHLIEDSTLVKGI--GIDGSVKLFHDYGSIKDV 209
DB 86 QICVGR-CLITIOISCKKVPDVLRSFLEDQITTFGVNNSQDKLERH-HQIDWRL 143
QY 210 EDLSLANOKIGGDKKWLASLETIVL-----CKELLKPNRIRLGNWEEYPLSKOOLYA 264
DB 144 VHRHLYHDL-----LSSSEETIVKYLGHGVTKDKELCMNMGARSLSHQIVQA 196
QY 265 ATDAY 269
DB 197 SHDY 201

RESULT 10
T19246
hypothetical protein C14A4.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T19246
R:McMurray, A.
submitted to the EMBL Data Library, June 1995
A:Reference number: Z19097
A:Accession: T19246
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-876 <ML>
A:Cross-references: EMBL:Z49909; PIDN:CAA90108.1; GSPDB:GN00020; CESP:C14A4.4
C:Experimental source: clone C14A4
C:Genetics:
A:Gene: CESP:C14A4.4
A:Map position: 2
A:Introns: 80/2; 557/2; 731/3; 833/1

Query Match 6.8%; Score 101; DB 2; Length 876;
Best Local Similarity 21.1%; Pred. No. 2;

Matches 64; Conservative 60; Mismatches 125; Indels 54; Gaps 15;

QY 15 ELAIDAIAASYNFSRSSSSAAP-----TVQATTSVGHEDPNQIPNNIR 64
DB 183 QMWVLEKPKTYNI-----SSDNSQAPFSSKLYVKKHAIKRGIVLHDDDESG-----RR 233
QY 65 QLEPRTSS-----TSYKREPLSRCARNPAMPFGRIYKSTATEVDKRAMOLIKY 117
DB 234 DMSIAETEEHEHPYIAETLHFVPEAQJSAECKFTALMKDPLMIDTKKELALYTK 293
QY 118 LDTKDESGIAFVGLDIEWRSPFRKGVLPKVAAYQICV-DSNY-CDVMHIF-HSGIPDS 174

DB 294 LNSVKE-----FAVDLEHQ--MRSYL-GLTCLIOISTFDEDFIDPPIMHVG--- 341
QY 175 LQHLIEDSTLVKKGIGIDGSVKLFHDYGSIKDVEDLSLANOKIGGDKKWLASLETIEN 234
DB 342 LNEFPANPRILKFKHSDSDSVLWLQDYGVHVNLPD--YVANKKL-KYPKESLAYL--T 397
QY 235 LVCKEILKPNRIRLGNWEEYPLSKOOLYAATDAVASWMLYKVL-----KDLPDAY 285
DB 398 LRFADVLDKQYGLADWBRARPLNNAIINVARREDTHLYSDMLRQOLKQDTKDLANY 457
QY 286 SGS 288
DB 458 SES 460

RESULT 11
E72674
hypothetical protein APE0818 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C:Accession: E72674
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Ta
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: E72674
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-445 <RAW>
A:Cross-references: DDBJ:AP000060; NID:95104188; PIDN:BA79797.1; PID:d1043583; PID:g
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE0818

Query Match 6.5%; Score 97.5; DB 2; Length 445;
Best Local Similarity 22.5%; Pred. No. 1.6;
Matches 65; Conservative 36; Mismatches 83; Indels 105; Gaps 15;

QY 33 SSSSSAAPT-----QATTSVHGE-----EDPNQIPNNRRLPRSTTSYKRRPPLS 82
DB 5 SKNSTAPSSSHAGESTADVLGEYEVLPKDYPOEARKRLMLNLERGVKYSTLG 64
QY 83 RCRARNPAMPFGRIYKSTATEVDKRAMQILKVDTRKDESGIAFV-----GT 132
DB 65 VSNRA-YFYQMRRLR-----PISDSILERLELATODDLAGIFPAAYVDYQKVG 115
QY 133 DI-----EMRPSFRKGVLPKGVATVQICVDNSNTCDVMHIFHSGIPSOHLIEDSTL 184
DB 116 DVDRLVRLVAEMARA-----NPASAKVFLDS-----LSALETRL----- 149
QY 185 VKYIGIGIDGSVKLFHDYGSIKDVEDLSLANOKI-GGDKKWLASLETIVCKELKP 243
DB 150 -----GLVGAIR-----VSEHNVREFESYLEARRSGMDMGCTGA----- 185
QY 244 NRIR-----LGNMEFYPLSKOOL-----QYATDAVASWMLYKVL 279
DB 186 DRLRYRLMALEEG-----YVLTKQALRGLIRRYQASQPGVAHYIKSLK 230

RESULT 12
F97039
hypothetical protein CAC1132 [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C:Accession: F97039
R:Rolling, J.; Breton, G.; Omeilenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; L
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: F97039

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 10:16:45 ; Search time 10 Seconds
(without alignments)

1194.518 Million cell updates/sec

Title: US-09-896-186B-24

Perfect score: 1491
Sequence: 1 MSSNMWIDAFTEBELLAID.....YASWHLKYVLKLDPAVSGS 288

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	352	23.6	1432	1 WRN_HUMAN	Q14191 homo sapien
2	339	22.7	1401	1 WRN_MOUSE	O09053 mus musculu
3	136.5	9.2	910	1 Y068_CAEEL	P34607 caenorhabdi
4	95	6.4	508	1 CPV1_RAT	P22443 rattus norv
5	93.5	6.3	1525	1 POLG_STREVM	P09732 s genome po
6	93	6.2	2224	1 PA5_HUMAN	P12259 homo sapien
7	93	6.2	3079	1 IRA2_YEAST	P19158 saccharomyc
8	92	6.2	758	1 LEU2_SCHPO	O14289 schizosacch
9	91.5	6.1	473	1 VL2_HPV03	P36744 human papil
10	91	6.1	503	1 CPV1_MOUSE	P28649 mus musculu
11	89.5	6.0	930	1 DPOL_HAETN	P43741 haemophilus
12	89.5	6.0	1324	1 MSH6_NKATH	O04716 arabidopsis
13	89.5	6.0	1487	1 BLA_DROME	O09418 drosophila
14	88	5.9	269	1 NK2B_BRARE	O09481 brachydania
15	88	5.9	416	1 PROA_VIBCH	O09481 vibrio chol
16	88	5.9	784	1 Y063_CAEEL	P34603 caenorhabdi
17	87	5.8	579	1 MTC1_BACST	P43423 bacillus st
18	87	5.8	3137	1 CA36_CHICK	P15983 gallus gall
19	86.5	5.8	346	1 PESTS_ECOLI	P06128 escherichia
20	86.5	5.8	470	1 PLSB_CUCSA	O39639 cucumis sat
21	86	5.8	552	1 FMR2_APLCA	P08021 alypsia cal
22	85.5	5.7	329	1 IDI2_STREY	O94095 streptococc
23	85.5	5.7	1394	1 E75B_DROME	P16772 drosophila
24	85.5	5.7	2298	1 C005_HUMAN	O09373 homo sapien
25	85.5	5.7	480	1 D6DH_SOYBN	O06558 glycine max
26	84.5	5.7	830	1 YASB_CLOAB	P94605 clostridium
27	84.5	5.6	777	1 YASB_SCHPO	O10146 schizosacch
28	83.5	5.6	237	1 TRPC_THEVO	O09799 thermoplasm
29	83	5.6	1129	1 PRXB_SOLFU	P34094 solanum tub
30	83	5.6	229	1 RPE_CHLUP	O94820 chlamydia p
31	82	5.5	492	1 P6GD_SCHPO	P76812 schizosacch
32	82	5.5	540	1 CH60_THERR	O60024 thermomane

ALIGNMENTS

RESULT 1	ID	WRN_HUMAN	STANDARD:	PRT: 1432 AA.
AC	Q14191:			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Werner syndrome helicase.			
GN	WRN OR RECQL2 OR RECQ3.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=9618115; PubMed=8602509;			
RA	Yu C.-E., Oshima J., Fu Y.-H., Wajsmann E.M., Hisama F., Allisch R.,			
RA	Matthews S., Nakura J., Miki T., Oualis S., Martin G.M., Mulligan J.,			
RA	Schellenberg G.D.;			
RT	"Positional cloning of the Werner's syndrome gene.";			
RL	Science 272:258-262(1996).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Paepker B.W., Gayle M., Brady W., Swartz A., Gillett L.A., Allisch R.S.,			
RA	Mulligan J., Galas D., Fu Y.-H.;			
RT	"Genomic structure of the human Werner's gene and cloning of the			
RL	mouse homolog.";			
RN	Submitted (Aug-1999) to the EMBL/Genbank/DBJ databases.			
RN	[3]			
RP	SUBCELLULAR LOCATION.			
RX	MEDLINE=98284027; PubMed=9618508;			
RA	Marchiaki R.A., Lombard D.B., Johnson F.B., Guarente L.;			
RT	"Nucleolar localization of the Werner syndrome protein in human			
RL	cells.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 95:6887-6892(1998).			
RN	[4]			
RP	REPEATS.			
RX	MEDLINE=99160561; PubMed=10049920;			
RA	Kusano K., Berres M.E., Engels W.R.;			
RT	"Evolution of the RECQ family of helicases: a Drosophila homolog,			
RL	Dmblm, is similar to the human Bloom syndrome gene.";			
RN	Genetics 151:1027-1039(1999).			
RN	[5]			
RP	REVIEW ON VARIANTS.			
RX	MEDLINE=99235545; PubMed=10220139;			
RA	Moser M.J., Oshima J., Monnat R.J. Jr.;			
RT	"WRN mutations in Werner syndrome.";			
RL	Hum. Mutat. 13:271-279(1999).			
RN	[6]			
RP	VARIANT ARG-1367.			
RX	MEDLINE=9021029;			
RA	Ye L., Miki T., Nakura J., Oshima J., Kamino K., Rakugi H.,			
RA	Ikegami H., Higaki J., Edland S.D., Martin G.M., Ogihara T.;			
RT	"Association of a polymorphic variant of the Werner helicase gene with			
RL	myocardial infarction in a Japanese population.";			
RL	Am. J. Med. Genet. 68:494-498(1997).			

34	81.5	5.5	462	1	LEU2_LISIN	Q92a26 listeria in
35	81.5	5.5	636	1	GYRB_THEMA	P77993 thermotoga
36	81.5	5.5	1002	1	YEMA_DROME	P25992 drosophila
37	81	5.4	309	1	CC2B_ARATH	P25859 arabidopsis
38	81	5.4	459	1	Y819_PYROO	O58549 pyrococcus
39	81	5.4	560	1	VAOX_PENSI	P56216 penicillium
40	81	5.4	864	1	CHEA_BORBU	Q44737 borrelia bu
41	81	5.4	885	1	PMC2_HUMAN	O01780 homo sapien
42	80.5	5.4	399	1	RPP_P14HA	P22044 human parat
43	80.5	5.4	451	1	NEF1_MOUSE	O92113 mus musculu
44	80.5	5.4	473	1	VL2_HPV28	P50799 human papil
45	80.5	5.4	543	1	PROA_LEGPN	P21347 legionella


```

RX MEDLINE-98284027; PubMed=9618508;
RA Marcinjak R.A., Lombard D.B., Johnson F.B., Guarente L.;
RT "Nucleolar localization of the Werner syndrome protein in human
RN cells."
RL Proc. Natl. Acad. Sci. U.S.A. 95:6887-6892(1998).
RP SEQUENCE FROM N.A.
RA Paeppe B.W., Gayle M., Brady W., Swartz A., Gillett L.A., Alisch R.S.,
RA Mulligan J., Gais D., Fu Y.-H.;
RT "Genomic structure of the human Werner's gene and cloning of its mouse
RT homolog."
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MAY BE INVOLVED IN THE CONTROL OF GENOMIC STABILITY.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE HELICASE FAMILY. RECD SUBFAMILY.
CC -!- SIMILARITY: CONTRAINS 1 HRDC DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D86527; BAA20270.1; -
DR EMBL; D86526; BAA20269.1; -
DR EMBL; AF091215; AAC78077.1; -
DR MGD; MGI:109635; Wtn.
DR InterPro; IPR002562; 3_5_exonuclease.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR002121; HRDC.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004589; RecQ.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; helicase_C; 1.
DR Pfam; PF00570; HRDC; 1.
DR Pfam; PF01612; 3_5_exonuclease; 1.
DR SMART; SM00474; 35EXOC; 1.
DR SMART; SM00487; DEXOC; 1.
DR SMART; SM00490; HELICC; 1.
DR SMART; SM00341; HRDC; 1.
DR TIGRfams; TIGR00614; recq; 1.
KW Helicase; ATP-binding; Nuclear protein.
FT NP_BIND 535 542 ATP (BY SIMILARITY).
FT SITE 632 635 DEAD BOX.
FT DOMAIN 1115 1194 HRDC.
FT DOMAIN 1387 1390 POLY-SER.
FT CONFLICT 101 101 N -> S (IN REF. 3).
FT CONFLICT 228 228 V -> A (IN REF. 3).
FT CONFLICT 250 250 M -> S (IN REF. 3).
FT CONFLICT 452 452 L -> V (IN REF. 3).
FT CONFLICT 459 459 K -> T (IN REF. 3).
FT CONFLICT 468 468 C -> R (IN REF. 3).
FT CONFLICT 619 619 Q -> K (IN REF. 3).
FT CONFLICT 800 800 Q -> R (IN REF. 3).
FT CONFLICT 1021 1021 L -> S (IN REF. 3).
FT CONFLICT 1145 1145 A -> T (IN REF. 3).
FT CONFLICT 1181 1182 VG -> LE (IN REF. 3).
FT CONFLICT 1252 1252 V -> A (IN REF. 3).
FT CONFLICT 1308 1308 I -> L (IN REF. 3).
FT CONFLICT 1356 1356 V -> A (IN REF. 3).
SQ SEQUENCE 1401 AA; 157256 MW; 949060992467FB8C CRC64;

Query Match 22.7%; Score 339; DB 1; Length 1401;
Best Local Similarity 36.6%; Pred. No. 1.5e-20;
Matches 74; Conservative 42; Mismatches 72; Indels 14; Gaps 5;

```

```

DB 87 KRSRAVADLCSEKNCYLFHTSSMSVFFQGLKMLENKSIRKAGVIGDQWKLRRDD 146
OY 204 VSIRKQVEDISDIANOKIGGDKRWGLASLEFVLCHELLKPNRIKGNWFFPLSRQOLOY 263
DB 147 VKIESFVELDVANERKCAETWMSLNGLVKHYLGKQLKDKSIRCSNWSNPFLEDOXLY 206
OY 264 AATDAVASHMLKRYLKDLPDAV 265
DB 207 AATDAVAGLITVYKLGNGLDTV 228

RESULT 3
Y068_CAEEL
ID Y068_CAEEL STANDARD; PRT; 910 AA.
AC P34607;
DT 01-FEB-1994 (rel. 28, Created)
DT 01-FEB-1994 (rel. 28, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Hypothetical 105.6 kDa protein ZK1098.8 in chromosome III.
GN ZK1098.8.
OC Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxId=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2.
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA Sims M., Smaildon N., Smith A., Smith M., Sonhammer E., Staden K.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Wohldman P.;
RT *2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.
RL Nature 368:32-38(1994).
CC -!- SIMILARITY: NO RIBONUCLEASE D.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z22176; CAA80137.1; -
DR PIR; S40930; S40930.
DR WormPep; ZK1098.8; CE00370.
DR InterPro; IPR002562; 3_5_exonuclease.
DR Pfam; PF01612; 3_5_exonuclease; 1.
DR SMART; SM00474; 35EXOC; 1.
KW Hypothetical protein.
SQ SEQUENCE 910 AA; 105569 MW; 5512D15423517FCD CRC64;

Query Match 9.2%; Score 136.5; DB 1; Length 910;
Best Local Similarity 24.8%; Pred. No. 0.0011;
Matches 77; Conservative 34; Mismatches 110; Indels 89; Gaps 13;

```

```

QY 120 TKRDESLAEVGLDIEMRSPFRKGVLPKVAATVOI-----CYD-----SNYCDYMH1 166
D 422 SLSDPAPYVGVGDESEKPKPSNLFAVHDSKIAIIGLFKKNCVWLVDCEELKANMAD----- 477
QY 167 FHSGIPOSLOHLEDSTLVKVGIGICDD-----SVKLFHDYGSIKDVEDLSD 214
D 478 --DMWOKFASRLGDSFVKVGFDMRNDLDAMATIPALKSSMKI-----EDTKNAFPLKR 530
QY 215 LANOKIGD-----KKMGSLAFPELVCKELLKNRIRLGWMEFPLSKOOLQVAT 266
D 531 LAENVDIDMEIIELEPKRTKRLADLHYHGLGLDKTE--QCSNMOCRLPKKOIYVAL 588
QY 267 DAYASWHLXK 276
D 589 DAVVVEFEK 598

RESULT 4
CPVL_RAT STANDARD; PRT; 508 AA.
ID CPVL_RAT
AC P22443;
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Cytochrome P450 19 (Aromatase) (EC 1.14.14.1) (CYPXIX) (Estrogen synthetase) (P-450AROM).
GN CYP19 OR AROM.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=90220647; PubMed=2157976;
RA Hickey G.J., Krasnow J.S., Beattie W.G., Richards J.S.;
RT "Aromatase cytochrome P450 in rat ovarian granulosa cells before and after luteinization: adenosine 3',5'-monophosphate-dependent and independent regulation. Cloning and sequencing of rat aromatase cDNA and 5' genomic DNA."
RT Mol. Endocrinol. 4:3-12(1990).
RL -1- FUNCTION: CATALYZES THE FORMATION OF AROMATIC C18 ESTROGENS FROM C19 ANDROGENS.
CC -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) -> ROH + oxidized flavoprotein + H(2)O.
CC -1- SUBCELLULAR LOCATION: Membrane-bound.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M33986; AAA41044.1; -.
DR PIR: A36121.
DR InterPro: IPR001128; Cytochrome_P450.
DR Pfam: PF00067; P450.1.
DR PRINTS: PR00385; P450.
DR PROSITE: PS00086; CYTOCHROME_P450; 1.
KW Electron transport; Oxidoreductase; Monooxygenase; Membrane; Heme.
FT BINDING 437 437 HEME.
SQ SEQUENCE 508 AA; 58411 MW; C0ADPBF0FD80AB352 CRC64;

Query Match 6.4%; Score 95; DB 1; Length 508;
Best Local Similarity 23.1%; Pred. No. 1.6;
Matches 42; Conservative 25; Mismatches 41; Indels 74; Gaps 9;
QY 93 REGGILYKTAETVDKRAMQLIKVLDFKRDESGIAFVGIDIEW---RPSFRKGVLP-GK 148

```

```

D 115 RFG-----SKRGLOCIQM-----HENGILFNPNPSLMRTVRFPMKALGPGL 157
QY 149 VATVOICVD-----SNYCDYMH1FHSIGPOSLOHLED-STLVKVGIGID 192
D 158 IRMEVCEVESIKOHLDRLGVDVTSNGYVDVVTL-----MRHIMDTNSTLFLGLPLD 209
QY 193 GDSYK-----LFHDYGSIKDVED-----LSDLANOKIGD 223
D 210 ESSIVKIGGYFNAMQALLKPNIFEFISULYRKYERSVADLDELIELVEKKRQKVS 269
QY 224 KK 225
D 270 EK 271

RESULT 5
POLG_STEVM STANDARD; PRT; 1525 AA.
ID POLG_STEVM
AC P09732; Q88781; Q88782; Q88783; Q88784; Q88785; Q88786; Q88787; Q88788;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polypeptide [Contains: Capsid protein C (core protein); Matrix protein (Envelope protein M); Major envelope protein E; Nonstructural proteins NS1, NS2A, NS2B; Protease/helicase (EC 3.4.21.98) (NS3)] (Fragment).
DE St. Louis encephalitis virus (strain MS1-7).
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Flavivirus.
OX NCBI_TaxID=11081;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87122172; PubMed=3027980;
RA Trent D.W., Kinney R.M., Johnson B.J.B., Vorndam A.V., Grant J.A., Deibel V., Rice C.M., Hahn C.;
RT "Partial nucleotide sequence of St. Louis encephalitis virus RNA: structural proteins, NS1, NS2A, and NS2B."
RL Virology 156:293-304(1987).
CC -1- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
CC NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polypeptide, commonly with Asp or Glu in the p6 position. Cys or Thr in p1 and Ser or Ala in p1'.
CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M16614; AAA47786.1; -.
DR PIR: A27531; GNMV55.
DR InterPro: IPR000069; Flav1_M.
DR InterPro: IPR001157; Flav1_NS1.
DR InterPro: IPR000752; Flav1_NS2A.
DR InterPro: IPR000487; Flav1_NS2B.
DR InterPro: IPR001122; Flav1_capsidC.
DR InterPro: IPR000336; Flav1_glycoprote.
DR InterPro: IPR002535; Flav1_propep.
DR Pfam: PF00869; Flav1_glycoprot; 1.
DR Pfam: PF00948; Flav1_NS1; 1.
DR Pfam: PF01002; Flav1_NS2B; 1.
DR Pfam: PF01003; Flav1_capsid; 1.
DR Pfam: PF01004; Flav1_M; 1.

```



```

DR Pfam: PF01005; Flavi_N52A; 1.
DR Pfam: PF01570; Flavi_propep; 1.
DR Pfam: PF02832; Flavi_glycop.C; 1.
DR ProDom: PD001496; Flavi_N51; 1.
DR ProDom: PD001556; Flavi_glycoprot; 1.
KW Polypeptide; Glycoprotein; Core protein; Coat protein;
KW Envelope protein; Hydrolase; Helicase; ATP-binding; Transmembrane;
KW Nonstructural protein.
FT INIT_MET 1 1 REMOVED FROM CAPSID PROTEIN C BY THE
FT CHAIN 1 121 CELLULAR AMINOPEPTIDASE.
FT PROPEP 122 213 CAPSID PROTEIN C.
FT CHAIN 214 288 ENVELOPE GLYCOPROTEIN M.
FT CHAIN 289 789 MAJOR ENVELOPE PROTEIN E.
FT CHAIN 790 1203 NONSTRUCTURAL PROTEIN NS1.
FT CHAIN 1204 1368 NONSTRUCTURAL PROTEIN NS2A.
FT CHAIN 1369 1499 NONSTRUCTURAL PROTEIN NS2B.
FT CHAIN 1500 >1525 PROTEASE/HELICASE (NS3).
FT TRANSMEM 108 119 POTENTIAL.
FT TRANSMEM 253 268 POTENTIAL.
FT TRANSMEM 274 288 POTENTIAL.
FT TRANSMEM 751 762 POTENTIAL.
FT TRANSMEM 768 787 POTENTIAL.
FT TRANSMEM 1173 1188 POTENTIAL.
FT DISULFID 291 318 BY SIMILARITY.
FT DISULFID 348 404 BY SIMILARITY.
FT DISULFID 362 393 BY SIMILARITY.
FT DISULFID 380 409 BY SIMILARITY.
FT DISULFID 478 576 BY SIMILARITY.
FT DISULFID 593 624 BY SIMILARITY.
FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 269 269 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 919 919 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 964 964 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 996 996 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1189 1189 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT NON_TER 1525 1525
SQ SEQUENCE 1525 AA; 167891 MW; 1A373FE511159F CRC64;

Query Match 6.3%; Score 93.5; DB 1; Length 1525;
Best Local Similarity 22.3%; Pred. No. 9.4;
Matches 47; Conservative 34; Mismatches 85; Indels 45; Gaps 11;

QY 82 SRCRRNPFPMFGRGRIILSKRTAEVDKRAMQLIVLDTKRE---SGIAFGDLIE-WR 137
DB 764 ADRSISLTLAVAGSILFLAVTSVQDSGCC-----AISLQRRELKCGGIGFYNDYEKKW 818
QY 138 PSFRKGVLT--PGKVAITVOICVDSNICYDVK-----HIFHSGLPQSLQHLIEDSTLVKGI 189
DB 819 SDYKFFLPPTGCLAVIQAHAHANGYCGIRSTRLEHLMENIQRELNAIFEDNE-IDLSV 877
QY 190 GIDGDSVKLFHDYGVSIKDVEDLSLANQIGGDKRWGLASLTETLVCKELKPNRIRLG 249
DB 878 VVQEDP-KYKRRAPRLKKLEDELDT-----GMKKWGTLTVEP-----RLG 918
QY 250 NMEFY---PLSKQQLQVATDAYASMHLYKV 277
DB 919 NNTFVVDGPEPTEK-----CPTANRAMNSFKV 944

RESULT 6
FAS_HUMAN STANDARD: PRT; 2224 AA.
AC P12259; Q14285;
DR 01-OCT-1989 (Rel. 12; Created)
DR 01-FEB-1994 (Rel. 28; Last sequence update)
DR 15-JUN-2002 (Rel. 41; Last annotation update)
DE Coagulation factor V precursor (Activated protein C cofactor).
GN F5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

```

```

OX NCBI_TaxId-9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-9223268; PubMed-1567832;
RA Crisp L.D., Moore K.D., Kane W.H.;
RT "Structure of the gene for human coagulation factor V.";
RL Biochemistry 31:3777-3785(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-8726086; PubMed-3110773;
RA Jenny R.J., Pittman D.D., Toole J.J., Kriz R.W., Aldape R.A.,
RA Hewick R.M., Kaufman R.J., Mann K.G.;
RT "Complete cDNA and derived amino acid sequence of human factor V.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:4846-4850(1987).
RN [3]
RP SEQUENCE OF 1-1600 FROM N.A.
RX MEDLINE-88107560; PubMed-2827731;
RA Kane W.H., Ichinose A., Hagen F.S., Davie E.W.;
RT "Cloning of cDNAs coding for the heavy chain region and connecting
RT of internal repeats.";
RL Biochemistry 26:6508-6514(1987).
RN [4]
RP SEQUENCE OF 1188-1215 AND 1315-2224 FROM N.A.
RX MEDLINE-8631365; PubMed-3092220;
RA Kane W.H., Davie E.W.;
RT "Cloning of a cDNA coding for human factor V, a blood coagulation
RT factor homologous to factor VIII and ceruloplasmin.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:6800-6804(1986).
RN [5]
RP TISSUE SEQUENCE FROM N.A.
RX MEDLINE-93203619; PubMed-8454869;
RA Shen N.L.L., Fan S.-T., Pyati J., Graf R., Lapolla R.J.,
RA Edgington T.S.;
RT "The serine protease cofactor factor V is synthesized by
RT lymphocytes.";
RL J. Immunol. 150:2992-3001(1993).
RN [6]
RP SULFATION.
RX MEDLINE-94264012; PubMed-8204629;
RA Pittman D.D., Tomkinson K.N., Michnick D., Seligsohn U.,
RA Kaufman R.J.;
RT "Posttranslational sulfation of factor V is required for efficient
RT thrombin cleavage and activation and for full procoagulant activity.";
RL Biochemistry 33:6952-6959(1994).
RN [7]
RP SULFATION.
RX MEDLINE-90366699; PubMed-2168225;
RA Hottin G.L.;
RT "Sulfation of tyrosine residues in coagulation factor V.";
RL Blood 76:946-952(1990).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 2065-2224.
RX MEDLINE-20052169; PubMed-1058686;
RA Macedo-Ribeiro S., Bode W., Huber R., Quinn-Allen M.A., Kim S.W.,
RA Ortel T.L., Bourenkov G.P., Bartunik H.D., Stubbs M.T., Kane W.H.,
RA Fuentes-Prior P.;
RT "Crystal structures of the membrane-binding C2 domain of human
RT coagulation factor V.";
RL Nature 402:434-439(1999).
RN [9]
RP VARIANT APCR GLN-534.
RX MEDLINE-94217810; PubMed-8164741;
RA Bertina R.M., Koelseman B.P.C., Koster T., Rosendaal F.R.,
RA Dirven R.J., de Ronde H., van der Velden P.A., Reitsma P.H.;
RT "Mutation in blood coagulation factor V associated with resistance to
RT activated protein C.";
RL Nature 369:64-67(1994).
RN [10]
RP VARIANT APCR GLY-334, AND VARIANT LYS-513.
RX Chan W.P., Lee C.K., Kwong Y.L., Lam C.K., Liang R.;
RT "A novel mutation of factor V gene in Hong Kong Chinese.";

```


RESULT 7
 ID IRA2_YEAST STANDARD; PRT: 3079 AA.
 AC P19158; 008239;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Inhibitory regulator protein IRA2.
 GN IRA2 OR GIC4 OR CCS1 OR YOL081W OR CO0985.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90318397; PubMed=2164637;
 RA Tanaka K., Nakatoku M., Tamanoi F., Kaziro Y., Matsumoto K., Toh-E A.;
 RT "IRA2, a second gene of Saccharomyces cerevisiae that encodes a
 RT protein with a domain homologous to mammalian ras GTPase-activating
 RT protein.";
 RL Mol. Cell. Biol. 10:4303-4313(1990).
 RN [2]
 RP SEQUENCE OF 1-2423 FROM N.A.
 RC STRAIN=S288C / FY1679;
 RX MEDLINE=95208358; PubMed=7900427;
 RA Zumbstein E., Griffin H., Schweizer M.;
 RT "Sequence of a 10.27 kb segment on the left arm of chromosome XV from
 RT Saccharomyces cerevisiae includes part of the IRA2 gene and a
 RT putative new gene.";
 RL Yeast 10:1383-1387(1994).
 RN [3]
 RP SEQUENCE OF 1982-3079 FROM N.A.
 RX MEDLINE=97321807; PubMed=9178509;
 RA Tzermita M., Katsoulou C., Alexandraki D.;
 RT "Sequence analysis of a 33.2 kb segment from the left arm of yeast
 RT chromosome XV reveals eight known genes and ten new open reading
 RT frames including homologues of ABC transporters, inositol
 RT phosphatases and human expressed sequence tags.";
 RL Yeast 13:583-589(1997).
 RN [4]
 RP IDENTIFICATION OF CCS1 AS IRA2.
 RX MEDLINE=92405229; PubMed=1326414;
 RA Busseureau F., Dupont C.H., Boy-Marcotte E., Mallet L., Jacquet M.;
 RT "The CCS1 gene from Saccharomyces cerevisiae which is involved in
 RT mitochondrial functions is identified as IRA2 an attenuator of RAS1
 RT and RAS2 gene products.";
 RL Curr. Genet. 21:325-329(1992).
 CC -1- FUNCTION: INHIBITORY REGULATOR OF THE RAS-CYCLIC AMP PATHWAY.
 CC -1- STIMULATES THE GTPASE ACTIVITY OF RAS PROTEINS.
 CC -1- SIMILARITY: CONTAINS 1 RAS-GAP DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M33779; AAA34710.1; -;
 CC EMBL; X83121; CA58201.1; -;
 CC EMBL; X75449; CA53302.1; -;
 CC EMBL; Z74823; CA99093.1; -;
 CC PIR; S11190; RGRY12.
 CC SGD; S0005441; IRA2.
 CC InterPro; IPR001936; RASGAP.
 CC Pfam; PF00616; RASGAP; 1.
 CC SMART; SM00333; RASGAP; 1.
 CC PROSITE; PS00509; RAS_GTPASE_ACTIV_1; 1.
 CC PROSITE; PS50018; RAS_GTPASE_ACTIV_2; 1.
 KW DOMAIN 1701 1890 RAS-GAP.

FT DOMAIN 399 409 POLY-SER.
 FT DOMAIN 412 416 POLY-SER.
 FT DOMAIN 520 528 POLY-ALA.
 FT DOMAIN 2469 2472 POLY-LEU.
 FT CONFLICT 2317 2317 I -> K (IN REF. 3).
 SQ SEQUENCE 3079 AA; 351631 MW; 651EB2A2EEB479C0 CRC64;
 Query Match 6.2%; Score 93; DB 1; Length 3079;
 Best Local Similarity 19.3%; Pred. No. 26;
 Matches 57; Conservative 43; Mismatches 112; Indels 84; Gaps 10;
 QY 9 DAFTEELLADDAIEASYNFSNRSSSSAAPTVOATTSVHGHEEDPQINIRQLPR 68
 DB 1976 DDFYKTFLLIDVDVGLGQGPMEFSN-----EIPYIRHMDYDELVEFNRRHAF 2028
 QY 69 SITSTSY-----KRPFLSCRRANFPNPFGRGLYKATVEYDKRAMOLIK-- 116
 DB 2029 NIETSTAVSPVSEHSTSEGIPIITLTMSNF-----SDRHVIDITVAAYKFLQIY 2077
 QY 117 -----VDTFRDSSGIAFVGLDIEMRSPFRKGYLPKGVATVOICVDSNGCDVMI 166
 DB 2078 ARIWTKHCLLIIDCTEPDEG-----GLDMKRFSLVMGLLP-EVAP-KNIGCYINWET 2131
 QY 167 F-----HSGIPQSLQHLIEDSTLVKIGIDGDSVYKLFHDYGVSTIKDVED 211
 DB 2132 FMDNYGKCLDKRNVYVSSKIPHYFINSNDSDEGLM-SVGTIGQGLKVLQDINRVSJLHDTL 2190
 QY 212 LSDLANOKIGDCKKGLASLETTLVCKELKPNRIKLGWMEFPLSKOOLQYATD 267
 DB 2191 YDEKRRN-----FTFVSLKIGDIYQVYLHETPROYKIRD 2224

RESULT 8
 ID LEU2_SCHPO STANDARD; PRT: 758 AA.
 AC 014289;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 3-isopropylmalate dehydratase (EC 4.2.1.33) (Isopropylmalate
 DE isomerase) (Alpha-IPW isomerase) (IPMI).
 GN SPAC9E9.03.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltham T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Wolkaert G., Aert R., Kobben J., Grymoprez B.,
 RA Weltjens I., Vansluis E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lebrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Punelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaune V., Mottier S.,
 RA Galbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garçon A., Rhode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,


```

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA MEDLINE=91378248; PubMed=1897929;
RA Terushima M., Toda K., Kawamoto T., Kuribayashi I., Ogawa Y.,
RA Maeda T., Shizuta Y.;
RT "Isolation of a full-length cDNA encoding mouse aromatase P450.";
RL Arch. Biochem. Biophys. 285:231-237(1991).
CC -1- FUNCTION: CATALYZES THE FORMATION OF AROMATIC C18 ESTROGENS FROM
CC C19 ANDROGENS.
CC -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized oxaloprotein + H(2)O.
CC -1- SUBCELLULAR LOCATION: Membrane-bound.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D00659; BAA00551.1; -
CC PIR: S13912; S13912.
CC MGI: MGI:88587; Cyp19.
DR InterPro: IPR001128; Cytochrome_P450.
DR Pfam: PF00067; P450; 1.
DR PRINTS: PR00385; P450.
DR PROSITE: PS00086; CYTOCHROME_P450; 1.
KW Electron transport; Oxidoreductase; Monooxygenase; Membrane;
KW Heme.
FT BINDING 437 437 HEME (BY SIMILARITY).
SQ SEQUENCE 503 AA; 58015 MW; 737400389D1AEFF1 CRC64;

Query Match 6.1%; Score 91; DB 1; Length 503;
Best Local Similarity 23.2%; Pred. No. 3.5;
Matches 38; Conservative 23; Mismatches 33; Indels 70; Gaps 8;

OY 93 REGGILYKATVEVDKRAMOLIKVLDTRKDESGIAFVGLDIET--RPSFRKGYL-PGK 148
DB 115 RFG-----SKRGLOCTIGM-----HENGILFNNPNSIMRTIRPFKALGTPGL 157
OY 149 VATVQICVD-----SNYCDVMHIFHSIGPOSQHLIED-STLVKVGIGID 192
DB 158 VRMEVCEVSEIKQHLDRLEGVDTGTYGVVDTL-----MRHLTDSNMLFGLPILD 209
OY 193 GDSVK-----LPHDYGVSTIKDYED 211
DB 210 ESAIVKIKQGYFNAMQALLIKNPFIKISWLYKRYKERSVKDKLD 253

RESULT 11
DPO1_HAEIN STANDARD; PRT; 930 AA.
AC P43741.
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA polymerase I (EC 2.7.7.7) (POL I).
GN POLA OR HI0856.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RA MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA Mckenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,

```

```

RA Scott J.D., Shirley R., Liu L.-I., Glodex A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Sauder D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhman J.L., Geoghegan N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RL Science 269:496-512(1995).
CC -1- FUNCTION: IN ADDITION TO POLYMERASE ACTIVITY, THIS DNA POLYMERASE
CC EXHIBITS 3' TO 5' AND 5' TO 3' EXONUCLEASE ACTIVITY (BY
CC SIMILARITY).
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + (DNA)(N).
CC -1- SUBUNIT: SINGLE-CHAIN MONOMER WITH MULTIPLE FUNCTIONS.
CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-A FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U32767; AAC22515.1; -
CC HSP: P00582; 1KFS.
DR TIGR: H10856; -.
DR InterPro: IPR002562; 3_5-exonuclease.
DR InterPro: IPR002421; 5_3-exonuclease.
DR InterPro: IPR001098; DNA_pol.
DR InterPro: IPR002298; DNA_pol.
DR InterPro: IPR000513; Exo_N.I.
DR InterPro: IPR003584; HHH_2.
DR Pfam: PF00476; DNA_pol_A; 1.
DR Pfam: PF01367; 5_3-exonuclease; 1.
DR Pfam: PF01612; 3_5-exonuclease; 1.
DR Pfam: PF02739; 5_3-exonuc_N; 1.
DR PRINTS: PR00866; DNAPOL1.
DR SMART: SM00474; 3EXOC; 1.
DR SMART: SM00475; 53EXOC; 1.
DR SMART: SM00279; HNH2; 1.
DR SMART: SM00482; POLA; 1.
DR TIGRFRMS: TIGR00593; polA; 1.
DR PROSITE: PS00447; DNA_POLYMERASE_A; 1.
KW Transferase; DNA-directed DNA polymerase; DNA replication; DNA repair;
KW Hydrolyase; Exonuclease; DNA-binding; Complete proteome.
SQ SEQUENCE 930 AA; 103740 MW; 226654BB7CFE730B CRC64;

Query Match 6.0%; Score 89.5; DB 1; Length 930;
Best Local Similarity 20.7%; Pred. No. 11;
Matches 61; Conservative 42; Mismatches 93; Indels 99; Gaps 15;

OY 24 ASYNFSRSSSSSSAPVQATTSVGHEDPNQIPNNIRQLPRSTSSSYKRPPLSR 83
DB 280 ARTEFRKRLMEYVANGADSTTQTT-----EQP-----VAMNYKATFSQ 318
OY 84 CRRNFPNRRFGGRILYSTATEVD-----KRAMOLIKVLDTRKDE-----SGIA 128
DB 319 SAVEMTPKIQI-DREKYETLLQADLTRWIEKLNAAKLIAV-DTETDSLDVYSANLVGIS 376
OY 129 FVGLDIEMRPSFRKGYLPCKVATVQICVDSNCDVWHIHSIGPOSQGH-----LI 179
DB 377 F-----ALENGEAAYLPLOLD--YIDA-----PKYLESTALAAIKPIL 413
OY 180 EDSTLVKVGIGIDGVSVKLFHDYGVSIKVE-----DLSDLANQKIGD 223
DB 414 ENPNHRIKIQNLKPE-SIFARHGIELOVEFDMLSLTTLNSTRGRHNDLAKRYLG-- 470
OY 224 KRWGLASLTETLVCKELLKPNRIRLGNMEFPLSKOOLQYATTDVYASVHLYKVL 278
DB 471 -----HETIAFSLACKGKSQ-LTFNQDPL-EGATEYAAEDADVTKLQOAL 515

```



```

RESULT 14
NK2B_BRARE
ID NK2B_BRARE STANDARD: PRT: 269 AA.
AC 090481;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Homeobox protein NK-2 homolog B (Homeobox protein NKX-2.2).
GN NKX2.2 OR NK2.2 OR NK2.2.
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Osteichthyes; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_Taxid=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95324401; PubMed=7600991;
RA Barth K.A., Wilson S.W.;
RT "Expression of zebrafish nk2.2 is influenced by sonic hedgehog/vertebrate hedgehog-1 and demarcates a zone of neuronal differentiation in the embryonic forebrain."
RL Development 121:1755-1768(1995).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN A CONTINUOUS NARROW BAND OF CELLS ALONG A BOUNDARY ZONE DEMARCATING THE LOCATION AT WHICH TWO OF THE EARLIEST NUCLEI IN THE BRAIN DIFFERENTIATE.
CC -1- SIMILARITY: BELONGS TO THE NK-2 HOMEBOX FAMILY.
CC -1- SIMILARITY: CONTAINS 1 HOMEBOX DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X85977; CAAS9967.1; -
DR HSP: P22808; INK3.
DR ZFIN: ZDB-GENE-980526-403; nkx2.2.
DR InterPro: IPR001356; Homeobox.
DR Pfam: PF00046; homeobox; 1.
DR PRINTS: PR00024; HOMEBOX.
DR ProDom: PD000010; Homeobox; 1.
DR SMART: SM00389; HOX; 1.
DR PROSITE: PS00027; HOMEBOX_1; 1.
DR PROSITE: PS50071; HOMEBOX_2; 1.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein.
FT DNA_BIND 125 184 HOMEBOX.
SQ SEQUENCE 269 AA; 30306 MW; CF006285CD33D63 CRC64;

Query Match 5.9%; Score 88; DB 1; Length 269;
Best Local Similarity 22.2%; Pred. No. 2.7;
Matches 45; Conservative 32; Mismatches 64; Indels 62; Gaps 8;

QY 8 DDAFTTEELLAIIDAIEAS-VNFSRSSSSSSAAPTVOATTSVHGHEEDPQININIRQL 66
DB 69 DNYET-RWLTATDTSIOYSLHGLANSODTSAKSP-----EESADESNDKETS 116
QY 67 PRSITSTSTYKRPPLSRCRANRPFARFGRIIXSKAT-EVDKRAMQ----- 113
DB 117 SNCSDSGKKRRK-----RVLESNAQYTELERRFQOQRYLSAPEREHL 158
QY 114 --LTKVLDTR-----RDSGTAFVGLDIEMRSPFRKGLPGKVAIVQICVSNYC 161
DB 159 ASLIRLTPVOYKIMFQNHKKMKRARAREKMEVTHLPSRRVAVPVILVRDGKFCHTLKAQ 218
QY 162 DVNHIHFSGIP-----QSLQHL 178
DB 219 DLAATFOAGIGIFSAYSQSLQHM 241

```

```

RESULT 15
PROA_VIBCH
ID PROA_VIBCH STANDARD: PRT: 416 AA.
AC 09KPT9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gamma-glutamyl phosphate reductase (GPR) (EC 1.2.1.41) (Glutamate-5-semialdehyde dehydrogenase) (Glutamyl-gamma-semialdehyde dehydrogenase) (Gsa dehydrogenase).
GN PROA OR VC2273.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_Taxid=666;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gvinn M.L., Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A., Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D., Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P., McDonald L., Uitterlind T., Fleischmann R.D., Nierman W.C., White O., Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C., Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae."
RL Nature 406:477-483(2000).
CC -1- FUNCTION: CATALYZES THE NADPH DEPENDENT REDUCTION OF L-GAMMA-GLUTAMYL 5-PHOSPHATE INTO L-GLUTAMATE 5-SEMIALDEHYDE AND PHOSPHATE. THE PRODUCT SPONTANEOUSLY UNDERGOES CYCLIZATION TO FORM 1-PYRROLINE-5-CARBOXYLATE.
CC -1- CATALYTIC ACTIVITY: L-glutamate 5-semialdehyde + phosphate + NADP(+) = L-gamma-glutamyl 5-phosphate + NADPH.
CC -1- PATHWAY: Proline biosynthesis; second step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE GAMMA-GLUTAMYL PHOSPHATE REDUCTASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE004299; AAF95417.1; ALF_INT.
DR TIGR: VC2273; -
DR InterPro: IPR002086; Aldehyde_dehydr.
DR InterPro: IPR000965; Gglut_ph_reduct.
DR Pfam: PF00171; aldehyd; 1.
DR TIGRFAMs: TIGR00407; proA; 1.
DR PROSITE: PS01223; PROA; FALSE_NEG.
KW Oxidoreductase; Proline biosynthesis; NADP; Complete proteome.
SQ SEQUENCE 416 AA; 44504 MW; 7ACAS8E33735563 CRC64;

Query Match 5.9%; Score 88; DB 1; Length 416;
Best Local Similarity 21.7%; Pred. No. 4.8;
Matches 69; Conservative 54; Mismatches 109; Indels 86; Gaps 19;

QY 7 IDDAFTTEELLAIIDAIEASVNFSRSSSSSSAAPTVOATTSVHGHEEDPQININIRQL 63
DB 59 LSDAMLDRL-----NESRLQAIANDVNRVILKNDPV-GSEIDSRVLENGSLAR 108
QY 64 ROLPRSTSTSTYKRPP-----LSRC-RANRPFARFGRIIXSKATATVDKRAMQLIK 116
DB 109 RRVPLGVV-GVYIEARNVTIDIAALCLTKGNALILGKGETFEFSN-----MELVK 158
QY 117 VLDTKRDSGTAFVGLD-IEMRSPFRKGLPGKVAIVQICVSNYCDVMHIFSGIPQ-- 173
DB 159 VIOSALDKAGLPAASVQYIE-----KPDRELVTQLKMDYDMT-----IPRG 203

```



```

OY      174 SLGHLEDSLWVGIGIDSDSVKLFHDYGSIKDVEDLSL-----AOKIGGDCK 225
        | : ::::: || | | : : : | : : : | : : : | : : : | : : : |
Db      204 AGHKMCKENSTVPIIGI-----GFGSHLFEVDESAOLDSVAIVENAKV---QRP 251
        | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
OY      226 WGLASIFETVCKELKLKNRIRL-----GNMEFY--PLSKOOIQAA--TDAYA-----S 271
        :| | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db      252 SACNALDTLVHQAIAKPLDKLLAKNGKAFAEPAKAKALMSSAAAEIRDAQAAGDFRE 311
        :| | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
OY      272 WHLY-----KVLDLPDAP 285
        | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db      312 WLSTYLGVVVQDYOEAI 329

```

Search completed: June 6, 2003, 10:22:52
Job time : 12 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 10:19:55 ; Search time 32 Seconds
(without alignments)
1854.424 Million cell updates/sec

Title: US-09-896-186b-24

Perfect score: 1491
Sequence: 1 MSSSNWIDDAFTEELALDAIDAEASYNFSRSSSSSSAAPVQATTSVHGHEEDNPIN 60

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

SPRMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	1466.5	98.4	285	10 Q9TF68	Q9TF68 arabidopsis
2	1439.5	96.5	313	10 Q9SVW6	Q9SVW6 arabidopsis
3	339.5	22.8	1436	13 Q93530	Q93530 xenopus lae
4	339	22.7	988	11 Q92241	Q92241 mus musculu
5	339	22.7	1401	11 Q9JRD4	Q9JRD4 mus musculu
6	330.5	22.2	643	11 Q93948	Q93948 mus musculu
7	238	16.0	201	10 Q93VU9	Q93VU9 oryza sativ
8	229.5	15.4	346	5 Q9VE86	Q9VE86 drosophila
9	227.5	15.3	501	10 Q8S2J5	Q8S2J5 oryza sativ
10	221.5	14.9	353	5 Q961E1	Q961E1 drosophila
11	212.5	14.3	583	5 Q9VGN7	Q9VGN7 drosophila
12	179.5	12.0	582	10 Q9C7K6	Q9C7K6 arabidopsis
13	174.5	11.7	391	10 Q8SA18	Q8SA18 oryza sativ
14	173.5	11.6	494	10 Q93VS2	Q93VS2 oryza sativ
15	159.5	10.7	699	3 Q9C2I6	Q9C2I6 neurospora
16	158	10.6	123	10 Q22964	Q22964 arabidopsis

17	152.5	10.2	496	11 Q8VEG4	Q8VEG4 mus musculu
18	151.5	10.2	199	10 Q9LX79	Q9LX79 arabidopsis
19	151	10.1	496	4 Q9NVH0	Q9NVH0 homo sapien
20	142	9.5	599	4 Q96NPL	Q96NPL homo sapien
21	127.5	8.6	239	10 Q9SIR3	Q9SIR3 arabidopsis
22	121	8.1	505	10 Q9FIN8	Q9FIN8 arabidopsis
23	112.5	7.5	217	10 Q80886	Q80886 arabidopsis
24	111.5	7.5	210	10 Q9LHG3	Q9LHG3 arabidopsis
25	111	7.4	242	10 Q9LHG2	Q9LHG2 arabidopsis
26	111	7.4	300	2 Q9AL74	Q9AL74 citrobacter
27	110	7.4	238	10 Q9C7A5	Q9C7A5 arabidopsis
28	108	7.2	300	2 Q8RNV0	Q8RNV0 citrobacter
29	105	7.0	625	5 Q9VIF1	Q9VIF1 drosophila
30	103	6.9	220	10 Q9LHG1	Q9LHG1 arabidopsis
31	102.5	6.9	710	5 Q9NA69	Q9NA69 caenorhabdi
32	101	6.8	876	5 Q17951	Q17951 caenorhabdi
33	99.5	6.7	393	16 Q8R7C1	Q8R7C1 thermocaneer
34	98.5	6.6	265	10 Q9LHG5	Q9LHG5 arabidopsis
35	98.5	6.6	532	17 Q976Y3	Q976Y3 sulfolobus
36	98	6.6	216	10 Q8SA46	Q8SA46 hordeum vul
37	97.5	6.5	445	17 Q9YDV0	Q9YDV0 aeropyrum p
38	97	6.5	579	2 Q9R0K2	Q9R0K2 bacillus st
39	97	6.5	906	5 Q17281	Q17281 botryllus s
40	96.5	6.5	488	16 Q97JZ0	Q97JZ0 clostridium
41	95	6.4	455	10 Q9ZWS2	Q9ZWS2 vigna mungo
42	95	6.4	925	5 Q9GUG1	Q9GUG1 caenorhabdi
43	95	6.4	4589	5 Q76506	Q76506 tetrahymena
44	94.5	6.3	353	10 Q9LHG4	Q9LHG4 arabidopsis
45	94.5	6.3	416	5 Q96144	Q96144 plasmodium

ALIGNMENTS

RESULT 1

ID	Q9TF68	PRELIMINARY:	PRT:	285 AA.
AC	Q9TF68:			
DT	01-MAR-2001 (TREMBLrel. 16, Created)			
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	Exonuclease.			
GN	WRNEXO.			
OS	Arabidopsis thaliana (Mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsis.			
OX	NCBI_TaxID=3702;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=COLUMBIA; TISSUE=FLOWER.			
RA	Hartung F., Pichova H., Puchta H.;			
RT	"Molecular characterisation of RecQ homologues in Arabidopsis			
RT	thaliana.";			
RL	Nucleic Acids Res. 21:4275-4282(2000).			
DR	EMBL: AJ04476; CAC14871.1; ..			
DR	InterPro: IPR002562; 3_5_exonuclease.			
DR	Pfam: PF01612; 3_5_exonuclease; 1.			
DR	SMART: SM00474; 35EXOC; 1.			
SQ	SEQUENCE 285 AA; 31750 MW; A198CB93653E229B CRC64;			

Query Match 98.4%; Score 1466.5; DB 10; Length 285;
Best Local Similarity 99.0%; Pred. No. 6.7e-121;
Matches 285; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY	1	MSSSNWIDDAFTEELALDAIDAEASYNFSRSSSSSSAAPVQATTSVHGHEEDNPIN 60
DB	1	MSSSNWIDDAFTEELALDAIDAEASYNFSRSSSSSSAAPVQATTSVHGHEEDNPIN 60
QY	61	NIRQLPSTSTSTSYKRPPLSRCRRANFPAMFEGRLYSKTATVDKRAMQLIKVLDT 120
DB	61	NIRQLPSTSTSTSYKRPPLSRCRRANFPAMFEGRLYSKTATVDKRAMQLIKVLDT 120

Db 63 IRSSLL-----EEDV--LGFDEMPVYTKG-KTGKVALIQCVCSEKKCYLFHISP 110
 QY 168 HSGTQSOHLIEDSTLVKVGIGDSDSKYLFPHDGVSKSDVEDSLANOKIGDCKRG 227
 Db 111 MAGFEKGLRLEDESVKRVGVIGEDQWKLMSDELKLGFIELSEMANOKLRCKEWT 170
 QY 228 LASLTFEYVCKELLPNRIRLGNMEFYPLSKOOLQYATDAVASWHLYKVLKDL 281
 Db 171 ENGLKHLFKRQLYRKSKYRCNSMDIFLITEDQKLYATDAVAGLLITKYLEGM 224

RESULT 4

Q92241 PRELIMINARY: PRT: 988 AA.

AC 092241: 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE WRN protein.
 GN Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Paepker B.W., Gayle M., Brady W., Swartz A., Gillett L.A., Alisch R.S.,
 RA Mulligan J., Galas D., Fu Y.-H.;
 RT "Genomic structure of the human Werner's gene and cloning of the mouse
 RT homolog.";
 RT submitted (SEP-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF091216; AAC72359.1; -
 DR InterPro: IPR002562; 3_5_exonuclease.
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR001650; Helicase_C.
 DR pfam: PF01612; 3_5_exonuclease; 1.
 DR pfam: PF00270; DEAD; 1.
 DR pfam: PF00271; helicase_C; 1.
 DR SMART: SM00474; 35EXOC; 1.
 DR SMART: SM00487; DEXDC; 1.
 DR SMART: SM00490; HELIC; 1.
 DR TIGRFAMS: TIGR00614; recq; 1.
 KW ATP-binding; Helicase.
 SQ SEQUENCE 988 AA; 111513 MW; DBD4D77FA5505B2D CRC64;

Query Match 22.7%; Score 339; DB 11; Length 988;
 Best Local Similarity 36.6%; Pred. No. 5, 3e-21;
 Matches 74; Conservative 42; Mismatches 72; Indels 14; Gaps 5;

QY 88 NEPMRFGRIILYSTATATEVDKRAMQLIKVLDTRKDESGIAFVGLDIEMRPSFRKGVLP 147
 Db 37 NLPLEFPGSIYSEASD---CSFLSEDISMRLSDGV--VGFDMEWPPIYK----PG 86
 QY 148 K---VATVOICVDSNCDVMHIFHSGI--POSLOHLIEDSTLVKVGIGDSDVKLEHNDG 203
 Db 87 KRSRAVAVIQLCVSEKCYLFHISMSVFPQGLKMLENKSIRKAGVIGIGDQWKLKLRFD 146
 QY 204 VSIKVEDSLANOKIGDCKWGLASLTFEYVCKELLPNRIRLGNMEFYPLSKOOLQY 263
 Db 147 VKLESFVELTDVANENKLCALFTWSLNGLVKHYGLKQLDKSIRCSNWSNFFLITEDQKLY 206
 QY 264 AATDAVASWHLYKVLKDLPAV 285
 Db 207 AATDAVAGLLITKYLEGM 228

RESULT 5

Q9JUD4 PRELIMINARY: PRT: 1401 AA.

ID 09JUD4: 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE WRN protein.
 GN Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20221574; PubMed=10757812;
 RA Lombard D.B., Beard C., Johnson B., Marciniak R.A., Dausman J.,
 RA Bronson R., Buhlmann J.E., Lipman R., Curry R., Sharpe A.,
 RA Jaenisch R., Guarente L.;
 RT "Mutations in the WRN Gene in Mice Accelerate Mortality in a p53-Null
 RT Background.";
 RT Mol. Cell. Biol. 20:3286-3291(2000).
 DR EMBL: AF241636; AAF64490.1; -
 DR MGD: MGI:109635; Wrn.
 DR InterPro: IPR002562; 3_5_exonuclease.
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR001650; Helicase_C.
 DR InterPro: IPR002121; HRDC.
 DR pfam: PF01612; 3_5_exonuclease; 1.
 DR pfam: PF00270; DEAD; 1.
 DR pfam: PF00271; helicase_C; 1.
 DR SMART: SM00570; HRDC; 1.
 DR SMART: SM00474; 35EXOC; 1.
 DR SMART: SM00487; DEXDC; 1.
 DR SMART: SM00490; HELIC; 1.
 DR SMART: SM00341; HRDC; 1.
 DR TIGRFAMS: TIGR00614; recq; 1.
 KW ATP-binding; Helicase.
 SQ SEQUENCE 1401 AA; 157300 MW; 6CB330CD072C670 CRC64;

Query Match 22.7%; Score 339; DB 11; Length 1401;
 Best Local Similarity 36.6%; Pred. No. 8, 5e-21;
 Matches 74; Conservative 42; Mismatches 72; Indels 14; Gaps 5;

QY 88 NEPMRFGRIILYSTATATEVDKRAMQLIKVLDTRKDESGIAFVGLDIEMRPSFRKGVLP 147
 Db 37 NLPLEFPGSIYSEASD---CSFLSEDISMRLSDGV--VGFDMEWPPIYK----PG 86
 QY 148 K---VATVOICVDSNCDVMHIFHSGI--POSLOHLIEDSTLVKVGIGDSDVKLEHNDG 203
 Db 87 KRSRAVAVIQLCVSEKCYLFHISMSVFPQGLKMLENKSIRKAGVIGIGDQWKLKLRFD 146
 QY 204 VSIKVEDSLANOKIGDCKWGLASLTFEYVCKELLPNRIRLGNMEFYPLSKOOLQY 263
 Db 147 VKLESFVELTDVANENKLCALFTWSLNGLVKHYGLKQLDKSIRCSNWSNFFLITEDQKLY 206
 QY 264 AATDAVASWHLYKVLKDLPAV 285
 Db 207 AATDAVAGLLITKYLEGM 228

RESULT 6

O35948 PRELIMINARY: PRT: 643 AA.

AC 035948: 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Homolog of human Werner's syndrome protein.
 GN WRN OR WRN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wu J., He J., Mountz J.D.,
 RT "mouse WRN.";

```

RL Submitted (Apr-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U97045; AAB87366.1; -.
DR MGD: MGI:109635; Wtn.
DR Interpro: IPR002562; 3_5_exonuclease.
DR Interpro: IPR001410; DEAD.
DR Interpro: IPR001650; Helicase_C.
DR Interpro: IPR004589; RecQ.
DR Pfam: PF01612; 3_5_exonuclease; 1.
DR Pfam: PF00270; DEAD; 1.
DR SMART: SM00474; 35EXOC; 1.
DR SMART: SM00490; HELICC; 1.
DR TIGRFAMs: TIGR00614; recq; 1.
DR ATP-binding; Helicase.
SQ SEQUENCE 643 AA; 72821 MW; F9125C234ABB757 CRC64;

Query Match 22.28; Score 330.5; DB 11; Length 643;
Best Local Similarity 33.38; Pred. No. 1,7e-20;
Matches 75; Conservative 48; Mismatches 83; Indels 19; Gaps 6;

QY 61 NINROLP-----SITSTSYKRRPPLSRARNPAPRFGRLYKSTATEVDKRAMQLI 115
DB 5 SLGRKPEPMMSMQSORATEKACVQKNVEDNLPLEFGSITYSEAD-----CSPLS 60
QY 116 KVLDTKDESGIAFVGLDITMRSEFKGVLPK--VAVQICVDSNYCDVMHIFHSGL- 171
DB 61 EDISMRLSDGV--VGFDMEMPPLYK----PGKRRAVAIVQLCVSEKCYLFIHSSMSVF 114
QY 172 POSLQHLIEDSTLVKVGIGDGSVKLFHDYGSIKNVEPLSDLANOKISGDKWGLASL 231
DB 115 PGLKMLLENKSKIKKACVGEIGDQWMLRPFDYKLESEFVELTDVANKLCAETWSLNG 174
QY 232 TETLVCKELLKPNRIKLNFEFFPLSKQQLQYATDAYSMTLYK 276
DB 175 VKHVLGKQLLKDKSIRCSNMSNPFLTRDQLYATDAYAFTIYR 219

RESULT 7
Q93VU9 PRELIMINARY; PRT; 201 AA.
ID 093VU9;
AC 093VU9;
DT 01-DEC-2001 (TrEMBLrel. 19, Last Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE P0487H02.22 protein (P0682B08.3 protein).
GN P0487H02.22 OR P0682B08.3.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0487H02.22"
RL Submitted (Oct-2000) to the EMBL/GenBank/DBJ databases.
RN 12
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0682B08.3"
RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP002883; BAB67862.1; -.
DR EMBL: AP003578; BAB60933.1; -.
DR Interpro: IPR002562; 3_5_exonuclease.
DR Pfam: PF01612; 3_5_exonuclease; 1.
SQ SEQUENCE 201 AA; 18199 MW; 85F35C0CC5386265 CRC64;

Query Match 16.08; Score 238; DB 10; Length 201;
Best Local Similarity 31.08; Pred. No. 4.8e-13;

```

```

Matches 58; Conservative 36; Mismatches 87; Indels 6; Gaps 4;

QY 92 MREGRLIYKSTATEVDKRAMQLIKVLDTRKDESGIAFVGLDITMRSEFKGVLPKYAT 151
DB 10 VREGSATITDITVYSDVAADAMAGVRAARGNGL-IVGLDCEMRKN-HYSKRTSKVAV 67
QY 152 VOICVDSNYCDVMHIFHSGL-IPSLQHLIEDSTLVKVGIGDGSVKLFHDYGSINKDVE 210
DB 68 LQICAGRFCLVQLDFYANRPVAVADLLGDPSTVLVIGIGVEDPAKLEADYGVCAAPV 127
QY 211 DLSPLANOK---IGGDKKNGLASITETLVCKELLKPNRIKLNFEFFPLSKQQLQYATD 267
DB 128 DLEDACNRRRLGLVGTGRRLGLKGYAREVLNAMKPRPVMTSMKEKRELPADQVEYACID 187
QY 268 AYASMTL 274
DB 188 AYYSYKL 194

RESULT 8
Q9VE86 PRELIMINARY; PRT; 346 AA.
ID 09VE86;
AC 09VE86;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CG7670 protein.
GN CG7670.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Baller R.M., Basu P.V., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Paulis B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foaier C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalaal M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang X., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., Mcleod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclik J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinelt K., Remington K., Saunders R.D.C., Scheefel F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong X., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.W., Venter J.C.;

```

RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003721; AAF5541.1; -
DR FlyBase: FBgn0038608; CG670.
DR InterPro: IPR002562; 3_5_exonuclease.
DR Pfam: PF01612; 3_5_exonuclease; 1.
DR SMART: SM00474; 35EXOC; 1.
SQ SEQUENCE 346 AA; 39398 MW; 3150A9BAEFB74EB0 CRC64;

Query Match 15.4%; Score 229.5; DB 5; Length 346;
Best Local Similarity 26.9%; Pred. No. 5.5e-12;
Matches 66; Conservative 43; Mismatches 123; Indels 13; Gaps 5;

QY 41 TVQATTSVGHGHEEDNQIPNNIRQLPRSTSTSYKRPPLSRCARNEPAMFGGRILY 100
DB 68 TPQVTEKLAEMEENP---PKRRSSRLTSTRSMADGSPSPKEKPEKLPFKYGAIKY 124
QY 101 SKTATEVDKRAMQILKVLDTKRDESGIAFVGLDIEMRPSFRKGVLPKGVATVQICVDSNY 160
DB 125 FTESQDIAASADVDVLOWEKKODE--VYPMAFDMWPFSPQTG--PGKSAVQICVDEKC 180
QY 161 CDVWHIFH-SGIPSOHLIEDSTLVKVGIGIDGSAVKLFHDY-----GVSTKVEDLSD 214
DB 181 CYYQLTNVKKLPALVALINHPKRYLHGVMINKDFRRLARPEPVTAEPLIEKCVDLGL 240
QY 215 LANOKIGDKKMGKGLASTETLVCKELKPNRIRLGNWFEYPLSKQOLQYATDAYASWHL 274
DB 241 MCNEVCETGGRWLSLERLTNFIKAKMADSKRYMSKWHYIPLDENQMLYAAIDVYIGYI 300
QY 275 YKVLK 279
DB 301 YRELE 305

RESULT 9

Q882J5 PRELIMINARY; PRT; 501 AA.
AC Q882J5;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE OSJNB0021A09.10 protein.
GN OSJNB0021A09.10.
OS Oryza sativa (japonica cultivar-group).
CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
CC Euphorbiales; Oryzaeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa niponbare(ga3) genomic DNA, chromosome 1, BAC
clone:OSJNB0021A09.";
RL Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AP003218; BAB89458.1; -
SQ SEQUENCE 501 AA; 56404 MW; 153AB72F4064FA99 CRC64;

Query Match 15.3%; Score 227.5; DB 10; Length 501;
Best Local Similarity 35.6%; Pred. No. 1.4e-11;
Matches 58; Conservative 28; Mismatches 70; Indels 7; Gaps 5;

QY 114 LIKVLDTKRDESGIAFVGLDIEMRP-SFRKGVLPKGVATVQICVDSNYCDVWHIFHSG-I 171
DB 38 LDELIRLRRLNLTIVGLDVEWRPATYTHG--PGPVAVIDICV-GRCLIFQILHAYV 94
QY 172 POSIQLHIEDSTLVKVGIGIDGSAVKLFHDYGVSTKVEDLSDLANOKIGD--KKMGIA 229
DB 95 PDSLFDLADGRFTFVGIGHDDYAKLASHGHEVENVDRLYLAQITGKPAKLSAQI 154
QY 230 SLFETLVCKELKPNRIRLGNWFEYPLSKQOLQYATDAYASW 272
DB 155 GLVREVMGVNAPKPYHVVASAMDSWNLTPEQVMYACADAFASF 197

RESULT 10

Q961E1 PRELIMINARY; PRT; 353 AA.
AC Q961E1;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE GM14514P.
GN CG670.
OS *Drosophila melanogaster* (Fruit fly).
CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
CC Ephydroidae; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Gartin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nuno J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celisner S.;
RL Submitted (AUG-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AY051647; AAK93071.1; -
DR FlyBase: FBgn0038608; CG670.
DR InterPro: IPR002562; 3_5_exonuclease.
DR Pfam: PF01612; 3_5_exonuclease; 1.
SQ SEQUENCE 353 AA; 40248 MW; F9636CB263868AB4 CRC64;

Query Match 14.9%; Score 221.5; DB 5; Length 353;
Best Local Similarity 26.5%; Pred. No. 2.9e-11;
Matches 65; Conservative 43; Mismatches 124; Indels 13; Gaps 5;

QY 41 TVQATTSVGHGHEEDNQIPNNIRQLPRSTSTSYKRPPLSRCARNEPAMFGGRILY 100
DB 75 TPQVTEKLAEMEENP---PKRRSSRLTSTRSMADGSPSPKEKPEKLPFKYGAIKY 131
QY 101 SKTATEVDKRAMQILKVLDTKRDESGIAFVGLDIEMRPSFRKGVLPKGVATVQICVDSNY 160
DB 132 FTESQDIAASADVDVLOWEKKODE--VYPMAFDMWPFSPQTG--PGKSAVQICVDEKC 187
QY 161 CDVWHIFH-SGIPSOHLIEDSTLVKVGIGIDGSAVKLFHDY-----GVSTKVEDLSD 214
DB 188 CYYQLTNVKKLPALVALINHPKRYLHGVMINKDFRRLARPEPVTAEPLIEKCVDLGL 247
QY 215 LANOKIGDKKMGKGLASTETLVCKELKPNRIRLGNWFEYPLSKQOLQYATDAYASWHL 274
DB 248 MCNEVCETGGRWLSLERLTNFIKAKMADSKRYMSKWHYIPLDENQMLYAAIDVYIGYI 307
QY 275 YKVLK 279
DB 308 YRELE 312

RESULT 11

Q9VGN7 PRELIMINARY; PRT; 583 AA.
AC Q9VGN7;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE CG6744 protein (AT25352p).
GN CG6744.
OS *Drosophila melanogaster* (Fruit fly).
CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
CC Ephydroidae; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celnlker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sulten G.G., Morten J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazel R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Plannkuch C., Baldwin D.,
RA Ballaw R.M., Basu A., Bakendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Berkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson R., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fessler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merklow G., Mlshina N.V., Moberly C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheefel R., Smith H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Slier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Stryker R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
RN
RP
SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Drensek D., Farfan D., Flise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacle J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celnlker S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AEO03691; AAF54639.2; -;
DR EMBL: AY089433; AL90171.1; -;
DR FlyBase: FBgn0037901; CG6744.
DR InterPro: IPR002562; 3_5_exonuclease.
DR Pfam: PF01612; 3_5_exonuclease; 1.
DR SMART: SM00474; 35EXOC; 1.
SQ SEQUENCE 583 AA; 66866 MW; 562984B17952AD0 CRC64;

Query Match 14.3%; Score 212.5; DB 5; Length 583;
Best Local Similarity 33.5%; Pred. No. 5e-10;
Matches 61; Conservative 23; Mismatches 65; Indels 33; Gaps 6;
QY 109 KRAMOLIKVLDTRKDESGIAFVGLDIEF-----RPSFRKGVLPKRVATVOICVDSNYC 161
DB 71 KNCQCFKVL-----GFDCWITVGGSRP-----VALQSSHARGLC 108
QY 162 DVNHIFH-SGIPQSLQHLIEDSTLVVVGIGIDGDSYKLFHDYGVSIKDYEDSLDANQKI 220
DB 109 ALFRLCHMQIOPQDLKELLEDDSVIKVGAPOQEDAKLSHDYGVASTLDLRLFC--VM 166
QY 221 GGDKNKGLASLTETLVCKELKPNRIRLGNWMEFPYPSKQOOLQYAADAVASMLHY-KVKK 279
DB 167 AGHKPBGIGKLSKTHLNTYTLDKHMRILCSNWEAKTLEPKQLDYAANDALMAVAIYQKLCR 226
QY 280 DL 281
II

DB 227 DL 228
RESULT 12
Q9C7K6 PRELIMINARY; PRT; 582 AA.
AC Q9C7K6;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE 37-57, exonuclease, putative.
GN F1469.8.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN
RP
SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altieri H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn A., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gall J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hultzer L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremetska I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Mailli R., Marzilli A.,
RA Millscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Uiterback T., Van Aken S., Vaysberg M., Vysotskaya V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.M.,
RT "Sequence and analysis of chromosome 1 of the plant *Arabidopsis*
thaliana.";
RL Nature 408:816-820(2000).
DR EMBL: AC069159; AAG50917.1; -;
DR InterPro: IPR002562; 3_5_exonuclease.
DR InterPro: IPR002114; HPT_Serp-site.
DR Pfam: PF01612; 3_5_exonuclease; 1.
DR SMART: SM00474; 35EXOC; 1.
DR PROSITE: PS00589; PTS_HPR_SER; UNKNOWN_1.
KW Exonuclease.
SQ SEQUENCE 582 AA; 66467 MW; 34D752473F786680 CRC64;

Query Match 12.0%; Score 179.5; DB 10; Length 582;
Best Local Similarity 28.2%; Pred. No. 2.8e-07;
Matches 53; Conservative 36; Mismatches 66; Indels 33; Gaps 7;
QY 126 GIAFVGIDIMRPSFRKGVLPKRVATVOICVDSN--YCDVNHIFHS---IPQSLQHLIE 180
DB 377 GCRVVGIDCEKMKPKYIKGSKQNKVSIWQISDTRKIPFLDLIKLYNDASELINDCLSHITQ 436
QY 181 DSTLVKV-----GIGIDGDSYKLFHDYGVSIKDYEDSLDANQKI 220
DB 437 SKSTLKVSLTLEDYPRDKLSGYNFQCDIKQLSLSG-DLKCEPRYDMLDIGNVNEPF 495
QY 221 GGDKNKGLASLTETLVCKELKPNRIRLGNWMEFPYPSKQOOLQYAADAVASMLHYKVK 280
DB 496 G-----GIAGLTFKRIKLVSLNKR--RNSDEORPLSONQLEFYAALDAVLIHIFRVRD 548
QY 281 LPDAVSGS 288
DB 549 HPHDSSS 556
RESULT 13
Q8SA18 PRELIMINARY; PRT; 391 AA.
ID Q8SA18

AC Q8SA18;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE P0684E06.27 protein.
 GN P0684E06.27.
 OS Oryza sativa (Japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eubartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
 RT clone:P0684E06.27";
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL, AP003291; BAB85281.1; -
 SQ SEQUENCE 391 AA; 41510 MW; 9D5BD11CB8E05C2D CRC64;

Query Match 11.7%; Score 174.5; DB 10; Length 391;
 Best Local Similarity 28.0%; Pred. No. 4.5e-07;
 Matches 53; Conservative 32; Mismatches 89; Indels 15; Gaps 6;

OY 95 GGRILYKATPEVDRAMOLIKVLDTRKDESGIAFVGLDIEMRPSFRKGLPG----KV 149
 DB 194 GDTIRHITHTSHSLAQFINEIARERQGGI-IVGIDTEKRTD---LPMGKTCTYV 248
 OY 150 ATVOICVDNVCVNMHIFHSG--IPQSLQHLIEDSTLVKVGIGIDGSVKLFPHDYGVSIK 207
 DB 249 AVLOLCV-GRCRLLEFOYQAGNMVPHLEAFADPSVRFVAVNNDQRLANDCLRAVA 307
 OY 208 DYEDSDLANCKIGDK--KMGSLATELVCKELKPRIRLGNMEFFPLSKOOLQYTA 265
 DB 308 CAVDLRVAAAVLGGPELARAGLKRLLALTVMGAHMEKKNITKSRNGEPTLWEDVNVAC 367
 OY 266 TDAYASWHL 274
 DB 368 IDAYVSYEL 376

RESULT 14
 O93VS2 PRELIMINARY; PRT; 494 AA.
 ID O93VS2;
 AC O93VS2;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE P0487H02.21 protein (P0682B08.2 protein).
 GN P0487H02.21 OR P0682B08.2.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eubartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=4530;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
 RT clone:P0487H02.21";
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
 RT clone:P0682B08.2";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL, AP002883; BAB67861.1; -
 DR EMBL, AP003578; BAB60932.1; -
 DR InterPro; IPR002562; 3_5_exonuclease.

DR Pfam; PF01612; 3_5_exonuclease; 1.
 SQ SEQUENCE 494 AA; 51497 MW; 8F83F79405C2BD1E CRC64;

Query Match 11.6%; Score 173.5; DB 10; Length 494;
 Best Local Similarity 24.3%; Pred. No. 7.5e-07;
 Matches 69; Conservative 42; Mismatches 126; Indels 47; Gaps 11;

OY 13 EPEELADIDAEASYNSSSSSSSAAPVQATTSVHG-----HEEDN 56
 DB 204 DEELMG-----SAGRRARRPPSSSMGVGGEATETGIGGPAPPTSSSSPPPPPHRSPH 259
 OY 57 QIPNNIRQLPRSTSTSTSKRREPLSRCHARNEPAMRFGRIYKATPEVDRAMOLIK 116
 DB 260 HTPHPFRDRIDNSA-----PLLDQVDGGCTVSPFGSAI-DTIVSDAAADAEWR 309
 OY 117 -----VLDTRKDESGIAFVGLDIEMR--SFRKGLPGVAVYQICVDNVCVNMHIFHS 169
 DB 310 RVRASAAATPRGGGGL-IVGLDCEMRKPCDHLMPAVAP-TVALLQLCAGDS-CILQLLHV 366
 OY 170 G----IPQSLQHLIEDSTLVKVGIGIDGSVKLFPHDYGVSIKVEDSDLANCKIG---G 222
 DB 367 AGARRVPLVGLDLPADSVRLVIGIGENNAKLADGIVRCAAPVLEDCDRRLRLPG 426
 OY 223 DKRWGLASTLETIVCKELKPRIRLGNMEFFPLSKOOLQYAT 266
 DB 427 ARRLGLKGVREVILGLTMEKPMQVTRSDMERRLDAQVRYACS 470

RESULT 15

O9C216 PRELIMINARY; PRT; 699 AA.

ID O9C216;
 AC O9C216;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Related to WERNER syndrome heliase.
 GN 93G11.60.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 RN
 RP SEQUENCE FROM N.A.
 RA Schulte U., Algin V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
 RA Nyakatura G., Mewes H.W., Mannhaupt G.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 RN
 RP SEQUENCE FROM N.A.
 RA German Neurospora genome project;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL, AL513443; CAC28658.1; -
 DR InterPro; IPR002562; 3_5_exonuclease.
 DR Pfam; PF01612; 3_5_exonuclease; 1.
 DR SMART; SM00474; 35EXOC; 1.
 SQ SEQUENCE 699 AA; 77383 MW; 42FE96D430EAB1CA CRC64;

Query Match 10.7%; Score 159.5; DB 3; Length 699;
 Best Local Similarity 26.0%; Pred. No. 2.1e-05;
 Matches 52; Conservative 42; Mismatches 73; Indels 33; Gaps 9;

OY 97 RILYSTATEVDRAMOLIKVLDTRKDESGIAFVGLDIEMRPSFRKGLPG-VATVOIC 155
 DB 162 KVHYCSATTTTERVLYKQYF-MDEK-----ILGIDLEMEISAKSHSPRONVSYIQA 212
 OY 156 VDSNVCVNMHIFHSGI-POS-----LOHLIEDSTLVKVGIGIDGSVKLFPHDYGVSIK 207
 DB 213 SEKR-----IGFIHISLPRKDELASPLKQIIEDADVAVAGWIMGDCTRLKTFGIEMK 268
 OY 208 DYEDSDLANCKIGDK--KMGSLATELVCKELKPRIRLGNMEFFPLSKOOLQYTA 265
 DB 269 GYIELSHLKLKLVKYSASGEHKLVRRLVPLATLVKVEVLDLPMFK-GAVTSSMS-KPLNM 326
 OY 259 QOLQYAAATDAVASWHLKVL 278

Sat Jun 14 11:07:49 2003

us-09-896-186b-24.rspt

Page 8

Db 327 DQITSGSDAYAGVOLPMM 346

Search completed: June 6, 2003, 10:24:14
Job time : 35 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 10:22:40 ; Search time 15 seconds

(without alignments)
564.920 Million cell updates/sec

Title: US-09-896-186B-24

Perfect score: 1491

Sequence: 1 MSSSNWIDDAFTFEELALD.....YASWHLKYVLDPDAVSGS 288

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing:

Minimum Match 0%

Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	352	23.6	1432	US-08-781-891-71	Sequence 71, Appl
2	339	22.7	1401	US-09-127-670-6	Sequence 6, Appl
3	330	22.1	1401	US-08-781-891-206	Sequence 206, App
4	93	6.2	3079	PCT-US94-00198-4	Sequence 4, Appl
5	89	6.0	930	US-09-134-001C-5314	Sequence 5314, Ap
6	87.5	5.9	1375	US-09-722-139-2	Sequence 2, Appl
7	87.5	5.9	1375	US-09-721-832-2	Sequence 2, Appl
8	87.5	5.9	1375	US-09-721-832-2	Sequence 2, Appl
9	86.5	5.8	348	US-07-989-845-2	Sequence 2, Appl
10	86.5	5.8	348	PCT-US93-11298-2	Sequence 2, Appl
11	86.5	5.8	480	US-08-987-367-2	Sequence 2, Appl
12	82.5	5.5	480	US-08-987-367-4	Sequence 2, Appl
13	82	5.5	333	US-09-134-001C-4528	Sequence 4528, Ap
14	80	5.4	574	US-08-907-166-12	Sequence 12, Appl
15	79.5	5.3	431	US-08-845-258-34	Sequence 34, Appl
16	79.5	5.3	431	US-08-990-571-34	Sequence 34, Appl
17	79.5	5.3	431	US-08-990-571-34	Sequence 34, Appl
18	79.5	5.3	431	US-08-723-142A-34	Sequence 34, Appl
19	79.5	5.3	431	US-09-528-784A-34	Sequence 34, Appl
20	79	5.3	616	US-09-388-743-14	Sequence 14, Appl
21	78.5	5.3	927	US-08-895-601-6	Sequence 6, Appl
22	78.5	5.3	260	US-09-134-001C-5363	Sequence 5363, Ap
23	78.5	5.3	637	US-09-134-001C-5658	Sequence 5658, Ap
24	78.5	5.3	817	US-07-640-029-2	Sequence 2, Appl
25	78.5	5.3	822	US-07-997-133-1	Sequence 1, Appl
26	78.5	5.3	822	US-07-921-807B-4	Sequence 4, Appl
27	78.5	5.3	822	US-08-459-296-2	Sequence 2, Appl
27	78.5	5.3	822	US-08-441-944A-4	Sequence 4, Appl

28	78.5	5.3	822	US-08-451-822A-12	Sequence 12, Appl
29	78.5	5.3	822	US-08-439-992A-2	Sequence 2, Appl
30	78.5	5.3	822	US-08-323-430-12	Sequence 12, Appl
31	77.5	5.2	247	US-09-216-295-8	Sequence 8, Appl
32	76	5.1	1196	US-08-881-706-2	Sequence 2, Appl
33	75.5	5.1	407	US-09-271-438A-11	Sequence 11, Appl
34	75.5	5.1	701	US-09-132-028-2	Sequence 2, Appl
35	75.5	5.1	1056	US-08-687-289A-7	Sequence 7, Appl
36	75.5	5.1	1056	US-08-687-289A-8	Sequence 8, Appl
37	75.5	5.1	1058	US-08-687-289A-5	Sequence 5, Appl
38	75.5	5.1	1078	US-08-485-588-7	Sequence 7, Appl
39	75.5	5.1	1078	US-08-484-565-7	Sequence 7, Appl
40	75.5	5.1	1078	US-08-480-751-7	Sequence 7, Appl
41	75.5	5.1	1078	US-08-943-986-7	Sequence 7, Appl
42	75.5	5.1	1078	US-08-353-784-7	Sequence 7, Appl
43	75.5	5.1	1078	US-08-484-719B-7	Sequence 7, Appl
44	75.5	5.1	1078	US-08-484-159-7	Sequence 7, Appl
45	75.5	5.1	1088	US-08-485-588-6	Sequence 6, Appl

ALIGNMENTS

Result 1:
US-08-781-891-71
Sequence 71, Application US/08781891
Patent No. 6090620
GENERAL INFORMATION:
APPLICANT: Fu, Yang-Hui
APPLICANT: Yu, Chang-En
APPLICANT: Oshima, Junko
APPLICANT: Mulligan, John T.
APPLICANT: Schellengberg, Gerald D.
TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:
ADDRESSER: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,891
FILING DATE: 27-DEC-1996
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: No. 6090620Leuburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
FAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 71:
SEQUENCE CHARACTERISTICS:
LENGTH: 1432 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-781-891-71
Query Match 23.6%; Score 352; DB 3; Length 1432;
Best Local Similarity 37.7%; Pred. No. 8.3e-30;
Matches 75; Conservative 41; Mismatches 75; Indels 8; Gaps 3;
88 NPPARFGRIILYSTATVEVDRAMQLIKVLDTKRDESGIAFVGIDIEPRSGVYRG 147

[illegible]

```

RESULT 2
US-09-127-670-6
: Sequence 6, Application US/09127670
: Patent No. 6228583
: GENERAL INFORMATION:
: APPLICANT: Massachusetts Institute for Technology
: APPLICANT: Leonard P. Guarente
: APPLICANT: David A. Sinclair
: APPLICANT: David B. Lombard
: TITLE OF INVENTION: ASSAYS FOR COMPOUNDS WHICH EXTEND LIFE
: TITLE OF INVENTION: SPAN
: FILE REFERENCE: MIT-7720PA
: CURRENT APPLICATION NUMBER: US/09/127,670
: CURRENT FILING DATE: 1998-07-31
: EARLIER APPLICATION NUMBER: US 60/054,629
: EARLIER FILING DATE: 1997-08-04
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 6
: LENGTH: 1401
: TYPE: PRF
: ORGANISM: Murine
US-09-127-670-6

```

	Query Match	22.7%	Score 339	DB 4	Length 1401
	Best Local Similarity	36.6%	Pred. No. 2,2e-28		
	Matches	74	Conservative 42	Mismatches 72	Indels 14
				Gaps 5	
QY	88	NPPAMRGRIYLSKATATEVDRAMOLIKVLTDKRDESGIAFVGLDIEWRSPFKGVLP	147		
Db	37	NLEPFEPPGSIYVSYEASD---CSFLSEDISMRLSDDV--VGFDMEMPITYK----	PG 86		
QY	148	K---VAPVOICVDSNVCVDVMIHFHSGI--FOSLOHIEBSTLVKVGIGIDGDSVTLFFHDYG	203		
Db	87	KRSRAVAYIQLCVSENKCYLFPHISSMSVFPQGLMKLLENKSTIKKGVGIEGQMKLLDFD	146		
QY	204	VSIKDVEDSLDANOKIRIGDKRWGLASLFTETVYKELLKPNRIIRLGNNFEPFLSKOOLQY	263		
Db	147	VKLESEFELTDVANERKLCAGETWMSLNGVAKHYLGQQLKDKSIRCSNMSNPNFLLEDQKLY	206		
QY	264	AATDAAYASMHLYKVLKDLDPDAY	265		
Db	207	AATDAAYAGLITTYOKLGNLGDVY	228		

RESULT 3
 US-08-781-891-206
 : Sequence 206, Application US/08781891
 : Patent No. 6090620
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Fu, Ying-Hui
 : APPLICANT: Yu, Chang-Bn
 : APPLICANT: Oshima, Junko
 : APPLICANT: Mulligan, John T.
 : APPLICANT: Schellenberg, Gerald D.
 :
 : TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
 :
 : TITLE OF INVENTION: WERNER'S SYNDROME
 :
 : NUMBER OF SEQUENCES: 209

1 CORRESPONDENCE ADDRESS:
2 ADDRESSEE: SEED and BERRY LLP
3 STREET: 6300 Columbia Center, 701 Fifth Avenue
4 CITY: Seattle
5 STATE: Washington
6 COUNTRY: USA
7 ZIP: 98104-7092
8
9 COMPUTER READABLE FORM:
10 MEDIUM TYPE: Floppy disk
11 COMPUTER: IBM PC compatible
12 OPERATING SYSTEM: PC-DOS/MS-DOS
13 SOFTWARE: Patent In Release #1.0, Version #1.30
14 CURRENT APPLICATION DATA:
15 APPLICATION NUMBER: US/08/781,891
16 FILING DATE: 27-DEC-1996
17 CLASSIFICATION: 800
18 ATTORNEY/AGENT INFORMATION:
19 NAME: No. 6090620tenburg Ph.D., Carol
20 REGISTRATION NUMBER: 39,317
21 REFERENCE/DOCKET NUMBER: 240052.419
22 TELECOMMUNICATION INFORMATION:
23 TELEPHONE: (206) 622-4900
24 TELEFAX: (206) 682-6031
25 INFORMATION FOR SEQ ID NO: 206:
26 SEQUENCE CHARACTERISTICS:
27 LENGTH: 1401 amino acids
28 TYPE: amino acid
29 TOPOLOGY: linear
30 MOLECULE TYPE: protein
31 US-08-781-891-206

	Query Match	22.1%	Score 330;	DB 3;	Length 1401;	
	Best Local Similarity	36.0%;	Pred. No. 2.3e-27;			
Matches	72;	Conservative	43;	Mismatches	71;	Indels 14; Gaps 5
OY	88	NFPAMREGGRILLYSTATLDEVDKRAMQLIKVLDTRKDESGIAFVGDIIEWRSPFRKGVLPG	147			
Db	37	NLPLEFGGVSYSTEASD---CSFLSEDDISMRLSDGV--VGFDMEPPYTK---PG	86			
OY	148	K---VAFAQICVDNSNYCDVMHIFHSGI--POSTLOHLIEDSTLVKVGIGIDGDSVKLFHDYG	203			
Db	87	KRSRAAVIQGLCVASESKCYLFIHTSSMSYVPQGAKMLLENKSIIKAAGVLEGDOCKLLRPDD	146			
OY	204	VSIKDVEDLSLDANOKIGGDKKWGLASTLETLYCKELLKNRIRLGNNEFYPLSKOOLY	263			
Db	147	VKLSEFYVALTPVANEMKELCAEFTWSLNGJLVKHYLGQLLKDKSIRSNNMNPFLTEDOKLY	206			
OY	264	AATDAVASMWLYKYVLKDLPD	283			
Db	207	AATDAVASMWLYKYVLKDLPD	226			

```

1  RESULT 4
2  PCT-US94-00198-4
3  : Sequence 4, Application PC/TUS94/00198
4  : GENERAL INFORMATION:
5  : APPLICANT: Schering Corp.
6  : TITLE OF INVENTION: Ras Associated GAP Proteins
7  : NUMBER OF SEQUENCES: 6
8  : CORRESPONDENCE ADDRESS:
9  : ADDRESSEE: Schering Corp.
10 : STREET: 1 Girald Farms
11 : CITY: Madison
12 : STATE: New Jersey
13 : COUNTRY: USA
14 : ZIP: 94304-1104
15 : COMPUTER READABLE FORM:
16 : MEDIUM TYPE: Floppy disk
17 : COMPUTER: Macintosh
18 : OPERATING SYSTEM: 6 0 8
19 : SOFTWARE: Microsoft Word 5.1a
20 : CURRENT APPLICATION DATA:
21 : APPLICATION NUMBER: PCT/US94/00198

```

FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/004,824
FILING DATE: 15-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Lunn, Paul G.
REGISTRATION NUMBER: 32,743
REFERENCE/DOCKET NUMBER: DX0352 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (201)822-7255
TELEFAX: (201)822-7039
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 3079 amino acids
TYPE: amino acid
STRADEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Saccharomyces cerevisiae
PCT-US94-00198-4

Query Match 6.2%; Score 93; DB 5; Length 3079;
Best Local Similarity 19.3%; Pred. No. 2;
Matches 57; Conservative 43; Mismatches 112; Indels 84; Gaps 10;

QY 9 DAFTEELLDAIDAEIENFSSSSSSAAPTVOATTSVGHEDDPNOIPNNIRQLPR 68
DB 1976 DDFKTTPLIDVDVGLGQPKMERFN-----EIPYIREHMDYELFEFNHAFR 2028
QY 69 SITSTSY-----KREPLSCRARNFPMRFGGRILYSKATVCDKRAMOLIK- 116
DB 2029 NIEFSTAISPVHSTSESGIPITLINSN-----SDRHVIDVVAKFLDIY 2077
QY 117 -----VLDTRKDESGIAFYVGLDIEMRPFERKGVLPKGVATVQICVDSNYCDVAHI 166
DB 2078 ARWTKHCLIIDCEFEDEG---GLDKRKFLSLVMGLLP-EVAP-KKICIGCYRNVNET 2131
QY 167 F-----HSGIQSLOHLIEDSTLVKVGIGIDGSVKLFHDYGVSIKDYED 211
DB 2132 FMDNYGKCLDKDMNVYSSKIPHYFINSNSDEGLMK-SVGITGQGLKVLQDIRVSLHDITL 2190
QY 212 LSDLANOKIGGDKKMGSLAFETLVCKELKPNRIRLGNWEPYPSKQOLQYAAND 267
DB 2191 YDEKRNK-----FTPVSLKIDYIFQVLEHETPROKIKD 2224

RESULT 5
US-09-134-001C-5314
Sequence 5314, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134, 001C
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 5314
LENGTH: 930
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5314

Query Match 6.0%; Score 89; DB 4; Length 930;
Best Local Similarity 23.6%; Pred. No. 0.84;
Matches 71; Conservative 45; Mismatches 117; Indels 68; Gaps 16;

QY 1 MSSNWIIDAFTEELLDAID---ATEASYNFSSSSSSAAPTVOATTSVGHEDDPNO 57
DB 150 VESSNSMDTQAOQPSHTTINSASITQSDNEBNSRVSPANSKLTIESNTE-SNKEENTIE 208
QY 58 IPNNIRQLPRSTSS-TSYKRF-----PLSCRARNFPMR 93
DB 209 QPNKVRD---SITSPSSYKNIDKISNOBELNLPINEYENKRPJSTTSAG--PSSK 263
QY 94 FGGRLYSKATVENDKRAMOLIKVLDTR-----RDESGI--AFVGLDIEMRPSFR--KGV 144
DB 264 ---RVTVQLAEEGGSVNNHLIKYDOSITEGYDSDGIIKHADEANLIVYTEFVDDKV 320
QY 145 LPGAATVQICDSNYCVDMHFNHSGIFQSIQHLIEDSTLVKVGIGIDGSVK---LFH 200
DB 321 KSGDTMIVNIDKNTVPSDITDSF--AIFK-----IKDSGELIATGTDTNKKQITTYFT 373
QY 201 DYGVSIKQVE-----DLSDLANO--KIGGDKKMGSLAFETLVCKELKPNRIRLG 249
DB 374 DYVDKYEINKAHLKLTYSIDKSKVPNNMTKLDVEYKTLASSVKNKITV-EYQKPRENRTA 432
QY 250 N 250
DB 433 N 433

RESULT 6
US-09-722-139-2
Sequence 2, Application US/09722139
Patent No. 6355471
GENERAL INFORMATION:
APPLICANT: Beraud, Christophe
APPLICANT: Freedman, Richard
TITLE OF INVENTION: No. 6355471el motor proteins and methods for
TITLE OF INVENTION: their use
FILE REFERENCE: 1055
CURRENT APPLICATION NUMBER: US/09/722,139
CURRENT FILING DATE: 2000-11-24
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 1375
TYPE: PRT
ORGANISM: Human
US-09-722-139-2

Query Match 5.9%; Score 87.5; DB 4; Length 1375;
Best Local Similarity 19.0%; Pred. No. 2.3;
Matches 58; Conservative 52; Mismatches 116; Indels 79; Gaps 13;

QY 13 EELLDAIDAEIENFSSSSSSAAPTVOATTSVGHEDDPNOIPNNIRQLPRSTTS 72
DB 935 KEEQLA-----QYQANANOLIKQATFEFTANIAHQEE-----KVRKKEKELIES 979
QY 73 STSYKRFPLSCRARNFPMRFGGRILYSKATVENDKRAMOLIKVLDTRKDESGIAFYGL 132
DB 980 REKQREALERLARH--LERRHSALQRHSTLGTETIEEROKIASLNSGSRQSGIQ-ASL 1036
QY 133 DIEMRPSFRK-----GVLPGAATVQICVDSNYCDVMH 165
DB 1037 EAF-QEALFKQOELEFYEIQOLKQIYEVDDVQKDHNHTLGGKVAASSLPPSAEKSHLYP 1095
QY 166 I-----FHSQIPQSIQHLIED-STLVKVGIGIDGSVKLFHDYGVSIKDYEDLSLANQI 220
DB 1096 LMDARINAYIEEEVQRRLODLHRVISEGCSSTADTMK-----DNEKLNQTI 1142
QY 221 GGDKKMGSLAFETLVCKELKPNRIRLGNWEPYPSKQOLQYAATDAYAASHMLKYLKD 280
DB 1143 QRLKLY---ELCRDLCLVLMPEPPAAACAN--HPLLQODLVOLSLD---WK-----TE 1187
QY 281 LPDAV 285
DB 1188 IPDLV 1192

	Matches	54;	Conservative	34;	Mismatches	73;	Indels	97;	Gaps	12;
QY	86	ARNPAMFGGRILSKTATEY-----DKRAM-----	-QLIKYLDTKRDSG	126						
Db	107	AVNIPGLKSGELVLDGTIGDIYLGIKKWDEDAIAKLNPGLKLPSSQINAAV--RAADSG	165							
QY	127	IAPV-----GLDIEMRPSFRKGV-----	-LEGVATVOICVDS	158						
Db	166	TSFVFETSYLAKVNEEMKNNVGGSTGVKMPGIGAGCKNGNDIAFAVQRLLGALGYVEYA---	222							
QY	159	NYCDDVMHIFHGSIQSOHLIEDSTLVAVVGIGIDGSDSKLRHDIVYSITKDVEDLS-----	213							
Db	223	-YAKONNNIAYTKL-----	-ISADGKPVSPTPEENFANNAKAQADWSKTEPAQ	264						
QY	214	DLANOKIGDGDKWGKGLASTLETFLVCEKELKPNR---IRLGHWEEFYPLSKO--QLOGYAATD	267							
Db	265	DLTGNOK--GEDAMPITSTFTFLIHKKDQKPRPGQTEVLAFPFMYATGTAKKANODDYAS--	320							
QY	268	AAYSAMHLKYALKDLPDAY	285							
Db	321	-----LPPSV	325							

```

OY      127  IAPV-----GUDIENRPSFKGV-----LPGRATATQICVDS 158
           |||
           ::|||::|||
           '
Db      166  TSFEFTSYLAKVAEENKNNVGTGSTYKMFIGLGGGANDCIAFVORLGAIGVEYA--- 222
           |||
           ::|||::|||
           '
OY      159  NYCDVMHIFHSIGIPQSIQHLIEDSTLYKVGIGIDDSVYKLFPHDYGVSIKVEDLS----- 213
           -
           |||
           ::|||
           '
Db      223  -YAKONNLAYTKL-----ISADGKPVSPTEENPANAAGADMSKTPAQ 264
           |||
           ::|||
           '
OY      214  DLANOKIGSKKMGKGLASLETJLYCKELKPNR-----IKLGMWEPFLPSKQ--QOQVATD 267
           |||
           ::|||
           '
Db      265  DLTNOK--GEDAPITSTFTILHDKOKRPEOGTEVLYKFLFWAYKTGAQKANDLYAS-- 320
           |||
           ::|||
           '
OY      268  AYASWMLYKVLKDLDPDVA 285
           |||
           ::|||
           '
Db      321  -----LPDSV 325
           |||
           ::|||
           '

```

RESULT 10
 PCT-US93-11298-2
 Sequence 2, Application PC/TUS9311298
 GENERAL INFORMATION:
 APPLICANT: Genentech, Inc.
 TITLE OF INVENTION: METHOD OF CONTROLLING POLYPEPTIDE PRODUCTION IN
 TITLE OF INVENTION: BACTERIA
 NUMBER OF SEQUENCES: 31
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Genentech, Inc.
 STREET: 460 Point San Bruno Blvd
 CITY: South San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94080
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: patin (Genentech)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US93/11298
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Hasak, Janet E.
 REGISTRATION NUMBER: 28,616
 REFERENCE/DOCKET NUMBER: 752
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415/225-1896
 TELEFAX: 415/952-9881
 TELEX: 910/371-7168
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 348 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 PCT-US93-11298-2

APPLICANT: Dhugga, Kanarpal S.
APPLICANT: Singletary, George William
APPLICANT: Saunders, Court
APPLICANT: Witcher, Derrick
APPLICANT: Bruce, Wesley B.
APPLICANT: Sims, Lynne
APPLICANT: Lu, Guhua
TITLE OF INVENTION: PLANT URIDINE DIPHOSPHATE-GLUCOSE
TITLE OF INVENTION: DEHYDROGENASE GENES, PROTEINS, AND USES THEREOF
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/987,367
FILING DATE: 10-DEC-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Jones, Phil B.C.
REGISTRATION NUMBER: 38,195
REFERENCE/DOCKET NUMBER: 750027.404
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 480 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-987-367-2

Query Match	5.8%	Score 86.5	DB 5	Length 348
Best Local Similarity	20.9%	Pred. No. 0.34		
Matches	54	Conservative	34	Mismatches 73; Indels 97; Gaps 12
QY	86	ARNPFAHFGGRILSKATVEY-----DKRAM-----	QLIKVDTKRDESG	126
Db	107	AVNIPGKSGELVLDGTGDIYGIKRIKWPDEALAKINPGIKLPSONIAV--RAADSG		165

```

Query Match      5 88; Score 86.5; DB 4, Length 480;
Best Local Similarity 22.18; Pred. No. 0.56;
Matches 43; Conservative 32; Mismatches 63; Indels 57; Gaps 9

OY      81 LSRARAPNPFAMFGRIILSTKATVEDKRAMQLIKVLDTRKDESGIAFVGLDIEMWPSF 140
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
db      58 VKQCRGN-----LFFS--ADVEKHA-----FADIIQFSVN---TPRK 91

```


Db 11 GIKKAIKRL-ENSPYLYDTETTGDRIRLVOIGDEENTYVIDYEIODIEPLKRLINERG 69
QY 220 -IGDGRKRWL-----ASLTFELVCKELKPNRIRL-----GN 250
Db 70 IVGNHLKFDLKLRYVGFPPSATFTPTMIASILGYERSLHNHYVNLIGYSMDKSYQSD 129
QY 251 WEFPPLSKQOLQYATDAYASWHLKVLKDL 281
Db 130 WGVSVLSDAOQLKYANDVYVLRLELFPKMRDM 160

RESULT 15
US-08-845-258-34

; Sequence 34, Application US/08845258
; Patent No. 6183976

; GENERAL INFORMATION:

; APPLICANT: Reed, Steven G.

; APPLICANT: Lodes, Michael J.

; APPLICANT: Houghton, Raymond

; APPLICANT: Sleath, Paul R.

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS

; TITLE OF INVENTION: AND TREATMENT OF B. MICROTI INFECTION

; NUMBER OF SEQUENCES: 53

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SEED AND BERRY

; STREET: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; COUNTRY: USA

; ZIP: 98104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/845, 258

; FILING DATE: 24-APR-1997

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Makl, David J.

; REGISTRATION NUMBER: 31,392

; REFERENCE/DOCKET NUMBER: 210121.426C1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206) 622-4900

; TELEFAX: (206) 682-6031

; INFORMATION FOR SEQ ID NO: 34:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 431 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; US-08-845-258-34

Query Match 5.3%; Score 79.5; DB 4; Length 431;

Best Local Similarity 22.4%; Pred. No.2.9; Mismatches 106; Indels 87; Gaps 17;

Matches 66; Conservative 35; Mismatches 106; Indels 87; Gaps 17;

QY 9 DAFTEBELLAIDAIEASYNFSRSSSSSSAAPTVOATTSVHGHEEDPNQIPNNIRQLPR 68
Db 33 DNFTLDNPSAYEILRVSYNSNEFQVS-----PQINNEMESSRPE 73
QY 69 S-----ITSSSYRKRFPLSRCARFPPAMRGGRILYSKTATEVDKRAMQLIKVLDTRKD 123
Db 74 SNITVVVHSDVIMRF--NCKNR-----KSLSTHSLTEND-----ILKFGRI 113
QY 124 ESGI--AFVGLDIERPSFRKG---VLPKVATVOICVDSNYCDVMH--IFSGIRPQSLQH 177
Db 114 ELVSKCITIMGAGITASDLNLKGLGIFSPDKQST--NVC---NYFEDMHESYHILDTQRASD 169
QY 178 LIEDS-----TLVKVG--IGIDGDSVKLFHDYGVSIKDVEDLSDLANQKIGGDKKW 226
Db 170 CVSDGDADIDISNDFMDVQODNINSVDADSETCMANSGVTYVNTTENVSNSEN-----F 221

QY 227 G-LASL--TELIVCKELKPNRIRLGNMEFYPL-----SKOOLQYATDAYASW 272
Db 222 GKLSLVSTTTPLCRICL-----CGESDPGFLVTPCCKCKSLNIVHLECLRTW 269

Search completed: June 6, 2003, 10:24:59
Job time : 17 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OW protein - protein search, using sw model

Run on: June 6, 2003, 10:24:21 ; Search time 21 Seconds
(without alignments)
1415.867 Million cell updates/sec

Title: US-09-896-186b-24
Perfect score: 1491
Sequence: 1 MSSSNWIDAFTEBELLAID.....YASWHLKYLKLDPAVSGS 288

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEM_PUB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCRT_NEM_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEM_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEM_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/PCRTUS_PUBCOMB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEM_PUB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEM_PUB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEM_PUB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	151	10.1	147	10 US-09-729-674-279	Sequence 279, App
2	93	6.2	2224	9 US-10-115-563-14	Sequence 14, Appl
3	89.5	6.0	930	10 US-09-815-242-11133	Sequence 11133, A
4	87.5	5.9	1885	10 US-09-920-346-2	Sequence 2, Appl1
5	87	5.8	910	9 US-10-174-363-22	Sequence 22, Appl1
6	86.5	5.8	346	10 US-09-741-669-380	Sequence 380, App
7	86.5	5.8	367	9 US-10-268-441-15	Sequence 15, Appl
8	86.5	5.8	373	9 US-10-268-441-14	Sequence 14, Appl
9	86.5	5.8	593	10 US-09-815-242-12306	Sequence 12306, A
10	86	5.8	1042	10 US-09-888-615-74	Sequence 74, Appl
11	85.5	5.7	1394	9 US-10-108-605-213	Sequence 213, App
12	83.5	5.6	319	9 US-09-738-626-6900	Sequence 6900, Ap
13	82.5	5.5	305	9 US-09-738-626-6005	Sequence 6005, Ap
14	81	5.4	389	9 US-10-174-363-50	Sequence 50, Appl
15	80	5.4	574	10 US-09-391-340-12	Sequence 12, Appl
16	80	5.4	574	10 US-09-948-369-12	Sequence 12, Appl
17	79.5	5.3	431	9 US-09-286-488-34	Sequence 34, Appl
18	79.5	5.3	431	10 US-09-737-178-34	Sequence 34, Appl
19	79	5.3	374	9 US-10-268-441-6	Sequence 6, Appl1

	20	79	5.3	733	9	US-10-097-534-16	Sequence 16, Appl
	21	79	5.3	927	9 <th>US-10-097-534-15</th> <th>Sequence 15, Appl</th>	US-10-097-534-15	Sequence 15, Appl
	22	78.5	5.3	764	10 <th>US-09-925-302-714</th> <th>Sequence 714, App</th>	US-09-925-302-714	Sequence 714, App
	23	78.5	5.3	7968	9 <th>US-10-077-130-5</th> <th>Sequence 5, Appl1</th>	US-10-077-130-5	Sequence 5, Appl1
	24	77	5.2	515	10 <th>US-09-801-368-384</th> <th>Sequence 384, App</th>	US-09-801-368-384	Sequence 384, App
	25	76.5	5.1	822	9 <th>US-09-757-415A-2</th> <th>Sequence 2, Appl1</th>	US-09-757-415A-2	Sequence 2, Appl1
	26	76	5.1	393	10 <th>US-09-815-242-11856</th> <th>Sequence 11856, A</th>	US-09-815-242-11856	Sequence 11856, A
	27	76	5.1	396	10 <th>US-09-864-761-43232</th> <th>Sequence 43232, A</th>	US-09-864-761-43232	Sequence 43232, A
	28	76	5.1	1196	9 <th>US-09-823-394-2</th> <th>Sequence 2, Appl1</th>	US-09-823-394-2	Sequence 2, Appl1
	29	75.5	5.1	1078	9 <th>US-10-125-792-28</th> <th>Sequence 28, Appl</th>	US-10-125-792-28	Sequence 28, Appl
	30	75.5	5.1	1078	9 <th>US-10-125-778-28</th> <th>Sequence 28, Appl</th>	US-10-125-778-28	Sequence 28, Appl
	31	75.5	5.1	1078	10 <th>US-09-727-205-2</th> <th>Sequence 2, Appl1</th>	US-09-727-205-2	Sequence 2, Appl1
	32	75.5	5.1	1078	12 <th>US-10-002-854-2</th> <th>Sequence 2, Appl1</th>	US-10-002-854-2	Sequence 2, Appl1
	33	75.5	5.1	3782	9 <th>US-09-860-846-4</th> <th>Sequence 4, Appl1</th>	US-09-860-846-4	Sequence 4, Appl1
	34	75.5	5.1	3782	9 <th>US-09-988-384B-4</th> <th>Sequence 4, Appl1</th>	US-09-988-384B-4	Sequence 4, Appl1
	35	75.5	5.1	3782	9 <th>US-09-836-821-4</th> <th>Sequence 4, Appl1</th>	US-09-836-821-4	Sequence 4, Appl1
	36	75.5	5.1	3782	10 <th>US-09-861-289-4</th> <th>Sequence 4, Appl1</th>	US-09-861-289-4	Sequence 4, Appl1
	37	75	5.0	1385	10 <th>US-09-738-363-2</th> <th>Sequence 2, Appl1</th>	US-09-738-363-2	Sequence 2, Appl1
	38	75	5.0	1864	9 <th>US-09-832-292-27</th> <th>Sequence 27, Appl1</th>	US-09-832-292-27	Sequence 27, Appl1
	39	74.5	5.0	742	10 <th>US-09-732-180-9</th> <th>Sequence 9, Appl1</th>	US-09-732-180-9	Sequence 9, Appl1
	40	74.5	5.0	2092	9 <th>US-10-147-026-12</th> <th>Sequence 12, Appl1</th>	US-10-147-026-12	Sequence 12, Appl1
	41	74	5.0	517	10 <th>US-09-964-277-21</th> <th>Sequence 21, Appl1</th>	US-09-964-277-21	Sequence 21, Appl1
	42	74	5.0	580	10 <th>US-09-815-242-4959</th> <th>Sequence 4959, Ap</th>	US-09-815-242-4959	Sequence 4959, Ap
	43	74	5.0	589	10 <th>US-09-815-242-10803</th> <th>Sequence 10803, A</th>	US-09-815-242-10803	Sequence 10803, A
	44	74	5.0	665	10 <th>US-09-816-494-2</th> <th>Sequence 2, Appl1</th>	US-09-816-494-2	Sequence 2, Appl1
	45	74	5.0	665	10 <th>US-09-964-277-2</th> <th>Sequence 2, Appl1</th>	US-09-964-277-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-09-729-674-279
Sequence 279, Application US/09729674
Patent No. US2001003935A1
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John M.
APPLICANT: Lavalie, Edward R.
APPLICANT: Collins-Racie, Lisa A.
APPLICANT: Evans, Cheryl
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Agostino, Michael J.
APPLICANT: Steininger II, Robert J.
APPLICANT: Spaulding, Vikki
APPLICANT: Wong, Gordon G.
APPLICANT: Clark, Hilary
APPLICANT: Fechtel, Kim
APPLICANT: Genetics Institute, Inc.
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
FILE REFERENCE: 6055-64X
CURRENT APPLICATION NUMBER: US/09/729,674
CURRENT FILING DATE: 2000-12-04
PRIOR APPLICATION NUMBER: 09/539,330
PRIOR FILING DATE: 2000-03-30
NUMBER OF SEQ ID NOS: 283
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 279
LENGTH: 147
TYPE: PRT
ORGANISM: Homo sapiens
US-09-729-674-279
Query Match
Best Local Similarity 35.0%; Pred. No. 1.7e-07;
Matches 41; Conservative 20; Mismatches 48; Indels 8; Gaps 2;
111 IPQSILHIEDSTLVKVGIGDGSVKLFPHDYGISIKVEDLSIA-----NQTGDKK 225
123 LPTLLDIDAGITLKVGVGSEDSKLLDYLGVLRGCLDLRYLAKMQRNLLCNG--- 79
226 WGLASLTETLVCKELKRNRIKGNWEYPILSKQQLQYAAANDAVASWHLVYVLDLP 282


```
US-09-920-346-2
; Sequence 2, Application US/09920346
; Patent No. US20020081658A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Roy A.J.
; TITLE OF INVENTION: 18610, A NOVEL HUMAN TRANSIENT RECEPTOR
; FILE REFERENCE: MNI-182
; CURRENT APPLICATION NUMBER: US/09/920,346
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: USSN 60/221,925
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1885
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-920-346-2

Query Match
Best Local Similarity 21.1%; Score 87.5; DB 10; Length 1885;
Matches 54; Conservative 47; Mismatches 90; Indels 65; Gaps 12;

QY 53 EDPNQPNNIR--QLPR---SITSSSTYKRP---LSRCRANFPAMRFGGRILYSKT 103
DB 1464 EQQNDVRNIMETEMKXENNTSENTLKRVSLSAGFTDCHTSTIP-----VHSKQ 1515
QY 104 ATEVDKRAMOLIKVLDTKDESGIAFVGLDIEMRPSFR---KGVLPKQVATVQICVDS 158
DB 1516 AKKISRPP---STETTHEVDSKAALIPWLDPRPNRMPSEEGTLNGITSPFKPAMDT 1571
QY 159 NYCDDVHIFHSIGIPSLQHLIEDSTLYKIGID-----GQVYLFHDYGSINDVE 211
DB 1572 NT-----YISA-----VERNNLMRLSQSIPPTPPRGSPVTVYR-----LEE 1609
QY 212 LS-DLANOKIGDDKKWGLASLTETVCKELLPNRIRLGNWEYPLSKOOLQYATDAYA 270
DB 1610 SSPNLLNMSMSQGLAKIEFLSKEMGGSLRAV-----KYQCTWSEHDILK 1660
QY 271 SWHLKVLKDLDPDAYS 286
DB 1661 SCHLYIKSFLPEVVN 1676

RESULT 5
US-10-174-363-22
; Sequence 22, Application US/10174363
; Publication No. US20030077623A1
; GENERAL INFORMATION:
; APPLICANT: Butler, Karlene H.
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Havell, Leslie T.
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Sakai, Hajime
; TITLE OF INVENTION: Polynucleotides And Polypeptides Involved In Post-Transcriptional
; FILE REFERENCE: BBI454 US NA
; CURRENT APPLICATION NUMBER: US/10/174,363
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 60/298,973
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 22
; LENGTH: 910
; TYPE: PRT
; ORGANISM: Zea mays
US-10-174-363-22

Query Match
Best Local Similarity 5.8%; Score 87; DB 9; Length 910;
Matches 66; Conservative 40; Mismatches 95; Indels 120; Gaps 16;
```

```
QY 7 IDDAFTEELLIDA-----TEASYNFRSSSSSSAAPVQATTSV-HGHEEDP 55
DB 601 VNDPYLNLMLKTNALGINSLLQYVAS-----SSIPHSQVPTIILGADVSHG----- 651
QY 56 NQIPNNIRQLPRTSITSSYKRPPL-SRCRANFPAMRFGGR-----LYSKTAEVD- 108
DB 652 ---PGQDR---PSVAAYVSSRQMPILSRRA---SVHQSAHLEMSLSLFKRGTDG 701
QY 109 -----KR-----AMOLIKVLDTRDSGIAFVGLDIEMRPSFR 141
DB 702 LRESLIDFTSSGKRPEHIIIFRDVSESOFTOVINILDOIIEACKFLDERKWSFKFT 761
QY 142 -----KGVLPKQVATVQICVDSNYCDVMHIFHSIGIPSLQ-----HLI 179
DB 762 VTYAQKNHHTKFEQTASPDNVLPGLVYDSSKVCAPKPF-DRYMCALHGMIGITRPTNHYVL 820
QY 180 EDSTLYKVGIGIDDSVKLF-----HDYGVSIKD---VEDL 212
DB 821 HDE-----IGFSADEMQEYVHSLSYVQKSTAIASVYAPVCYAHAAQAVSTFLRLHEM 874
QY 213 SDLANOKIGDDKKWGLASLTET 233
DB 875 SDASSSGCGGHTSAGSAPYE 895

RESULT 6
US-09-741-669-380
; Sequence 380, Application US/09741669
; Patent No. US20020022718A1
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; TITLE OF INVENTION: Genes Identified as Required for
; FILE REFERENCE: ELTRA.009A
; CURRENT APPLICATION NUMBER: US/09/741,669
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 60/173005
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 380
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-741-669-380

Query Match
Best Local Similarity 5.8%; Score 86.5; DB 10; Length 346;
Matches 54; Conservative 34; Mismatches 73; Indels 97; Gaps 12;

QY 86 ARNFPAMRFGGRILYSKTATEV-----DKRAM-----OLIKVLDTKRDESG 126
DB 107 AVAIPGLKSELVLDGKTIGLDIYLGKIKKWDDEAIAKINPLKLPSONIAV-RRDGS 165
QY 127 IAFV-----GLDIEMRPSFRKV-----LPGKVAATVQICVDS 158
DB 166 TSEVFTSYLAKVNEEMKNVGTSGYKWPILGGLGKNDGIAAIVQRLPGAIGVEYA--- 222
QY 159 NYCDDVHIFHSIGIPSLQHLIEDSTLYKIGIDGDSYKLFHDYGSIKDVELS----- 213
DB 223 -YAKQNNLAVYTKL-----ISADGRVPSPTTEENFANAAGAWMSTFNO 264
QY 214 DLANOKIGDDKKWGLASLTETVCKELLPNR-----IRLGNWEYPLSKO--OLOYATD 267
DB 265 DLTNOK--GEDAMPITSTFILHDKOKRPGDSTVYLKFPDAVKTKGAGQANDLDYAS--- 320
QY 268 AYASWHLVYKVLKDLDPDAV 285
DB 321 -----LPDSV 325
```

```
RESULT 7
US-10-268-441-15
; Sequence 15, Application US/10268441
; Publication No. US20030084475A1
; GENERAL INFORMATION:
; APPLICANT: Coughlan, Edgar B.
; APPLICANT: Coughlan, Sean J.
; APPLICANT: Helentjaris, Timothy George
; APPLICANT: Jung, Rudolf
; APPLICANT: Li, Chun Ping
; APPLICANT: Nichols, Scott
; APPLICANT: Ripp, Kevin
; APPLICANT: Zheng, Peizhong
; TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS AND PROTEINS AFFECTING STORAGE
; TITLE OF INVENTION: ORGANELLE
; TITLE OF INVENTION: FORMATION AND METHODS OF USE
; FILE REFERENCE: B1392 US NA
; CURRENT APPLICATION NUMBER: US/10/268,441
; CURRENT FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: US/09/672,607
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 60/157209
; PRIOR FILING DATE: 1999-09-30
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-268-441-15

Query Match
Best Local Similarity 31.18; Pred. No. 2;
Matches 19; Conservative 16; Mismatches 25; Indels 1; Gaps 1;

QY 176 QHLEIDSTLVKVGIGIDGSKLPHDYGVSI-KDVEDLSLANOKIGDKKGLASTLET 234
DB 55 EHLEINPTRGHVGSSGNDPSLSTYPLILAILKDLTVVAAEHFYGDKMNYIILTEA 114
QY 235 L 235
DB 115 M 115

RESULT 8
US-10-268-441-14
; Sequence 14, Application US/10268441
; Publication No. US20030084475A1
; GENERAL INFORMATION:
; APPLICANT: Cahnoun, Edgar B.
; APPLICANT: Coughlan, Sean J.
; APPLICANT: Helentjaris, Timothy George
; APPLICANT: Jung, Rudolf
; APPLICANT: Li, Chun Ping
; APPLICANT: Nichols, Scott
; APPLICANT: Ripp, Kevin
; APPLICANT: Zheng, Peizhong
; TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS AND PROTEINS AFFECTING STORAGE
; TITLE OF INVENTION: ORGANELLE
; TITLE OF INVENTION: FORMATION AND METHODS OF USE
; FILE REFERENCE: B1392 US NA
; CURRENT APPLICATION NUMBER: US/10/268,441
; CURRENT FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: US/09/672,607
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 60/157209
; PRIOR FILING DATE: 1999-09-30
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 14
; LENGTH: 373
; TYPE: PRT
```

```
; ORGANISM: Trilicium aestivum
US-10-268-441-14

Query Match
Best Local Similarity 32.18; Pred. No. 2.1;
Matches 34; Conservative 18; Mismatches 31; Indels 23; Gaps 7;

QY 144 VLPGKATVOICVDSNYCDVWHIFHSIGIPSL-QHLEIDSTLVKVGIGIDGDSV-----K 197
DB 28 ILPERFANSELAPEAVYVALD-----GVSSVNGHIEFTPT-----DGHSLAKSEGS 73
QY 198 LFHDYGVSI-KDVEDLSLANOK-IGDDKKWGLASTLET-L-VCKEL 240
DB 74 IPNALVYSILKDYEAVEVAQHFEVGDGRKMGFLAVAEVAKACVRL 119

RESULT 9
US-09-815-242-12306
; Sequence 12306, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12306
; LENGTH: 593
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12306

Query Match
Best Local Similarity 5.8%; Score 86.5; DB 10; Length 593;
Matches 55; Conservative 47; Mismatches 113; Indels 47; Gaps 11;

QY 2 SSSNWIDATTEELALDAI--EASTNFSRSSSSSSAAPTQATTSVGHEDPQIP 59
DB 150 SYONILSKVFTLPDQFFILATATATVEVOODIREKINIAOTDIOIKSTK----- 199
QY 60 NNTIRPOLPRSTTSYSTSKRPLSCBARNFPMRFGGRILYSTATFVDKRAMQLIKVLD 119
DB 200 ---RRNLIFRYNPFYQOKFTLLDYIKTHDDA-----GLIYCSTRKQVE---ELQFALE 247
QY 120 TKRDESGIAFVGADIEWRPSFRKGVLPK-----VATVOICVDSNYCDVWHIFHSIGIPSL 175
DB 248 SQKTESYIYHAGLSNKREREAQNDLFDRYKVVATNAFQMGIDKSNVRVIYHNMGD 307
QY 176 QHLEIDSTLVKVG-IGIDGSKLPHDYGVSI-KDVEDLSLANOKIGDKKGLASTLET 227
```

Db 308 ESYQVAG--RAGRGKSECILLFSESDIMLHEFTVVSQADD--DYKDKM--GEKLTK 361
QY 228 LALTEFLVCKE-----LAKPN 244
Db 362 MIQTKRKCLEAVIVHFEPN 383

RESULT 10

US-09-888-615-74
; Sequence 74, Application US/09888615
; Patent No. US20020064856A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: WHITE, DAVID
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: CHARYCZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 74
; LENGTH: 1042
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-615-74

Query Match 5.8%; Score 86; DB 10; Length 1042;
Best Local Similarity 23.5%; Pred. No. 10;
Matches 72; Conservative 42; Mismatches 113; Indels 80; Gaps 15;

QY 4 SMTIDDAFTEELAIIDAIEASYNFSRSSSSSSAAPVQATTSVHGHEEDPNQINNI 63
Db 199 SNLQNTWKAPALLLSLQEV--FASISSTDAFESVALASLV-----QHLPLOMT 249
QY 64 RQLRSTSTSTSYKRFPLSR--CRARNEPAMRFGGRILYKSTATEVDKRAMOLIKVLDTK 121
Db 250 TVLIRSLTDPBNVDASMTQALCRIMDW-----LSWPLAQHVDIVWVIALLK----- 295
QY 122 RDESGIAFYG-----LDIE-----WRPSFRKG---VLDGKAVTQICVDSNATCD 162
Db 296 ---GLAAVQKFTLIDVTLLKIELVFNRLMFPPLVRPALVLSHMLLSFQHSPEAFHLI 351
QY 163 VMHT---FHS---GIPOSLOHLIEDSTLVKVGIGIDDSVKLPHFDYGV-----SI 206
Db 352 VPHVYNLVHSEKNGDLSSTFLVQVLELI-----HGMTHYSGFPLUYEPILEAI 402
QY 207 KDV-----EDLSLANOKIGDKKMGSLASLETFLVCK-ELLKPNRIKLNMEFYPLSKQO 260
Db 403 KDEPKPEEKIKILINQSAMTSQNSLSLASCRLSGSEKTKTGILNGN-TGWNVSVIQ 461
QY 261 LQYATD 267
Db 462 ALFWATD 468

RESULT 11

US-10-108-605-213
; Sequence 213, Application US/10108605
; Patent No. US20020160934A1
; GENERAL INFORMATION:
; APPLICANT: Broadus, Julie
; APPLICANT: Stam, Lynn
; APPLICANT: Bacimann, Jane
; APPLICANT: Kandar, Kim
; TITLE OF INVENTION: NOCTEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE
; TITLE OF INVENTION: PROTEINS ESSENTIAL FOR LARVAL VIABILITY AND USES THEREOF

; FILE REFERENCE: 31133B
; CURRENT APPLICATION NUMBER: US/10/108,605
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 09/761,142
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/176,418
; PRIOR FILING DATE: 2000-01-14
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 213
; LENGTH: 1394
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-108-605-213

Query Match 5.7%; Score 85.5; DB 9; Length 1394;
Best Local Similarity 22.4%; Pred. No. 17;
Matches 80; Conservative 52; Mismatches 140; Indels 85; Gaps 17;

QY 8 DDAFTEELAIIDAIEASYNFSR-----SSSSSSAAPVQATTSVHGHEEDPNQINNI 62
Db 371 DDELCQHHORLDSOLNVLCOKEFKDLDTALSSSANTGRTPAVTANEDADGFFRSTI 430
QY 63 RQRL---PSITSTSYKRFPLSR-----ANFPAMRG-----GRILYS- 101
Db 431 QOKIYRCPCTNQQCSILIRINNRQCYRLKCKCIANGSROAVRGVPRKREKARILAM 490
QY 102 -----KTATEVDKRAMOLIKVLD-----TKRDESGIAFYGLDIEMRPSFRK 142
Db 491 QOSTQNRGQORALATELDOPRLAVALRAHLETCFTEKEKVSAMRORARDC---PSYSM 547
QY 143 GYL-----PGKVAIVQICVDSNCDVMH--TFHSGIGQSLOHLEDS--TLVKYIGID 192
Db 548 PTLACPLNPAPELOSEQEFSSORFAHVIRGVIDFAGMTIPFOLLTQDDKFTLLKAGL-PD 606
QY 193 GDSVLFPHDYGVSKIDV-----EDLSLANOKIGDKKMGIA-----SLTEP-- 234
Db 607 ALFVRLICMFSSINSIICLQNGVRRALQNGANARFLVYSTFPAERMNSMNLDAEI 666
QY 235 -LVCK-ELLKPNRIKLNMEFYPLSKQOQY--AAVDAYASHMLKRYKVLDPD 283
Db 667 GLFCAIVLITPDRPGLRMLLEIKMYSRILKGLQYIVAQNRDPDPEFLAKLETMPD 723

RESULT 12

US-09-738-626-6900
; Sequence 6900, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6900
; LENGTH: 319

```
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6900

Query Match
Best Local Similarity 17.0%; Score 83.5; DB 9; Length 319;
Matches 52; Conservative 46; Mismatches 88; Indels 119; Gaps 13;

QY 31 SSSSSSSAA-----PTVATTSVHGHEEDPNQIPNNIRQLPRSTTSSTYKRPFLSRC 84
DB 13 ATSTDVAVVGRLFTADSAATAGSDAPEGQFPRT----- 49
QY 85 RANPAPMRFGGRIYLSKATEYDKRAMQILKYLTPKRES-----GIAVGL-----DIEM 136
DB 50 -----VHNRGETTLEQOPORVY-VLDGGEIDYVLSTVTPYGIASPRDASS 95
QY 137 RPSFRKGVLT-----PGKVAATVQICVDNVCYDVMHI-----FHS 169
DB 96 QRAYLENOADYQVTTSELNFEAIAALKPDILLSKRLVDSYDQSLQIAFTVUSIRP 155
QY 170 GIPOSLOHLIEDSTLVKVGIGIDGDSVKLFHDYGSIKDVEDLSLANOKIGGDKRWGLA 229
DB 156 GEPWKK-----ENFLTLADALGLEKRAVEYLNEXQTHV-----DAVRETIDGSPESLTV 203
QY 230 SLFETLVCKELKPNRIKL-GNMEF-----YPLSKQQLQYA-----ATDAYA 270
DB 204 R-----FMPGRTRLGYNLSFGAILKDLGLSRPEIONIDDLAVEISPENITDANG 253
QY 271 SMHLY 275
DB 254 DMIFY 258

RESULT 13
US-09-738-626-6005
; Sequence 6005, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6005
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6005

Query Match
Best Local Similarity 5.5%; Score 82.5; DB 9; Length 305;
Matches 41; Conservative 29; Mismatches 62; Indels 55; Gaps 7;

QY 127 IAFVGDIMRPSFRGVLPGKVAATVQICVDNVCYDVMHIHSGIPOSLOHL-----E 180
DB 15 VSFVG-----RPYTGKSTLTNALVGEKIAITANO-----PETHRHPIRGILVHND 58
```

```
QY 181 DSTLVKVGIGIDGDSVKLFHDYGSIKDVEDLSL-----ANORIGGDKRWGLA----- 229
DB 59 NAQIIVDPFGLRPRFTLIGERLNEAVKDYADVDDLIGFVPAANEKIGPDRMILEAVRK 118
QY 230 -----SLFETLVCKELKPNRIKLGWMEYRPLSKQQLQYAATDAYASWML 274
DB 119 VSPKPTPLIGITRADSVSRDLVAQDMAVHELLGNSSEVVPVSTSGENET-----L 171
QY 275 KYVKDL 281
DB 172 IKVMTDL 178

RESULT 14
US-10-174-363-50
; Sequence 50, Application US/10174363
; Publication No. US20030077623A1
; GENERAL INFORMATION:
; APPLICANT: Butler, Karlene H.
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Sakai, Hajime
; TITLE OF INVENTION: Polynucleotides And Polypeptides Involved In Post-Transcriptio
; FILE REFERENCE: B1454 US NA
; CURRENT APPLICATION NUMBER: US/10/174,363
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 60/298,973
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 50
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-174-363-50

Query Match
Best Local Similarity 5.4%; Score 81; DB 9; Length 389;
Matches 36; Conservative 24; Mismatches 53; Indels 58; Gaps 6;

QY 113 QLIKVLDTKRDESGIAFVGIDIEMRPSFR-----KGYLPKQVATVQI 154
DB 212 OFQVQVINELEQIIEACKCDKWKPEKFTYIVAQKNHHTREFQTNSEPNVPGTVVDQV 271
QY 155 CVDNCDVWHIHSIGIPOSLO-----HLIEDSTLVKVGIGIDGDSVKLF----- 199
DB 272 CHPKRF-DFYWCANAGNIGTSRPTHYVHLDE-----IGFSGDELQEFVHSLSYVQRS 324
QY 200 -----HDYGSIKDVEDLSLANOKIGGDKRWGLASLTE 233
DB 325 TTAISVAAPTAIAVLAALAAQVGTFRK-FEDMSDTSSSGGGHTSAGSAPVPE 374

RESULT 15
US-09-391-340-12
; Sequence 12, Application US/09391340A
; Patent No. US20020013455A1
; GENERAL INFORMATION:
; APPLICANT: Callen, Walter
; APPLICANT: Mather, Eric
; TITLE OF INVENTION: ISOLATION AND IDENTIFICATION OF NOVEL POLYMERASES
; FILE REFERENCE: 09010/027001
; CURRENT APPLICATION NUMBER: US/09/391,340A
; CURRENT FILING DATE: 1999-09-07
; EARLIER APPLICATION NUMBER: US 08/907,166
; EARLIER FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 574
```


TYPE: PRT
ORGANISM: Aquifex pyrophilus
US-09-391-340-12

Query Match 5.48; Score 80; DB 10; Length 574;
Best Local Similarity 21.98; Pred. No. 18;
Matches 33; Conservative 30; Mismatches 48; Indels 40; Gaps 6;

QY	170	GIPOSLOHLEDSFLVKGIGIDDSVKLF-----HDYGV---SIKDVEDLSLANQK- 219
		: : : : : : : : : : : : : : : : : : :
Db	11	GIKKAIKRL-ENSPYLIDTETTGDRIRLVOIGDEENTYVIDLYEIODIEPLRKLINEKG 69
QY	220	-IGGDKRWGL-----ASLTETLVCKELKPNRIRL-----GN 250
		: : : : : : : : : : : : : : : : : : :
Db	70	IVGHNLKFDLKYLRGIFPSAFEDTMIASYLIGYERHSLNHIYSNLLGYSMDSKYQTS 129
QY	251	WEFPLSKQOLQYATDAVASMHLKYVKDL 281
Db	130	WGASVLSDAOIKYANDVIYLRLELFPKMRDM 160

Search completed: June 6, 2003, 10:31:31
Job time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 10:22:55 ; Search time 308 Seconds
(without alignments)
602.867 Million cell updates/sec

Title: US-09-896-186B-24

Perfect score: 1491

Sequence: 1 MSSSNWIDAFTEELLAID.....YASWHLKYKLDLPDAVSGS 288

Scoring table:

BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 10%

Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/1/paa/PCYUS.COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06.COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07.COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08.COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081.COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082.COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083.COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084.COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085.COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086.COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087.COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088.COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089.COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090.COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091.COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092.COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093.COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094.COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095.COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096.COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097.COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US098.COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US099.COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US100.COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US101.COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US102.COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US60.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1491	100.0	288	19	US-09-513-996A-67237
2	1491	100.0	288	22	US-09-896-186B-24
3	1491	100.0	298	19	US-09-513-996A-67236
4	1439.5	96.5	313	21	US-09-708-427-27138
5	1439.5	96.5	313	22	US-09-896-186B-2
6	1439.5	96.5	313	23	US-09-906-226-50

7	1433	96.1	288	19	US-09-513-996A-28856	Sequence 28856, A
8	1028	68.9	197	19	US-09-513-996A-67238	Sequence 67238, A
9	996	66.8	197	19	US-09-513-996A-28857	Sequence 28857, A
10	976.5	65.5	222	21	US-09-708-427-27139	Sequence 27139, A
11	896	60.1	177	19	US-09-513-996A-28858	Sequence 28858, A
12	876.5	58.8	202	21	US-09-708-427-27140	Sequence 27140, A
13	717	48.1	157	20	US-09-620-393B-6113	Sequence 6113, A
14	676.5	45.4	290	23	US-09-906-226-16	Sequence 16, App1
15	676.5	45.4	290	27	US-60-218-993-16	Sequence 16, App1
16	545.5	36.6	301	21	US-09-708-427-65911	Sequence 65911, A
17	545.5	36.6	301	21	US-09-708-427-83273	Sequence 83273, A
18	545.5	36.6	303	21	US-09-708-427-65910	Sequence 65910, A
19	545.5	36.6	303	21	US-09-708-427-83272	Sequence 83272, A
20	475	31.9	173	21	US-09-708-427-65912	Sequence 65912, A
21	475	31.9	173	21	US-09-708-427-83274	Sequence 83274, A
22	352	22.6	1409	23	US-09-948-941-397	Sequence 397, App
23	352	22.6	1432	9	US-08-594-242-71	Sequence 71, App1
24	352	22.6	1432	20	US-09-618-166-71	Sequence 71, App1
25	352	23.6	1432	21	US-09-791-537-602	Sequence 602, App
26	352	23.6	1432	22	US-09-896-186B-18	Sequence 18, App1
27	352	23.6	1432	23	US-09-948-941-273	Sequence 273, App
28	339.5	22.8	1436	21	US-09-791-537-106290	Sequence 106290, A
29	330.5	22.2	643	21	US-09-791-537-82396	Sequence 82396, A
30	330.5	22.2	643	23	US-09-906-226-49	Sequence 49, App1
31	330	22.1	1401	20	US-09-618-166-206	Sequence 206, App
32	330	22.1	1401	21	US-09-791-537-128052	Sequence 128052, A
33	270	18.1	216	23	US-09-906-226-38	Sequence 38, App1
34	254	17.0	234	23	US-09-906-226-14	Sequence 14, App1
35	254	17.0	234	27	US-60-218-993-14	Sequence 14, App1
36	249.5	16.7	210	26	US-10-219-999-33507	Sequence 33507, A
37	249.5	16.7	210	27	US-60-324-109-18761	Sequence 18761, A
38	249.5	16.7	213	23	US-09-906-226-10	Sequence 10, App1
39	249.5	16.7	213	27	US-60-218-993-10	Sequence 10, App1
40	245	16.4	236	26	US-10-219-999-60555	Sequence 60555, A
41	242.5	16.3	241	26	US-10-219-999-56100	Sequence 56100, A
42	238	16.0	197	23	US-09-906-226-12	Sequence 12, App1
43	238	16.0	197	27	US-60-218-993-12	Sequence 12, App1
44	238	16.0	307	27	US-60-243-468-753	Sequence 753, App
45	237.5	15.9	230	26	US-10-219-999-48774	Sequence 48774, A

ALIGNMENTS

RESULT 1
US-09-513-996A-67237
; Sequence 67237, Application US/09513996A
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-709P
; CURRENT APPLICATION NUMBER: US/09/513,996A
; CURRENT FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 81028
; SEQ ID NO 67237
; LENGTH: 288
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 1..288
; OTHER INFORMATION: any n or xaa = unknown
; OTHER INFORMATION: Location 1..288 / Ceres Seq. ID 2177599
US-09-513-996A-67237

Query Match 100.0% Score 1491; DB 19; Length 288;
Best Local Similarity 100.0%; Pred. No. 6.7e-154;
Matches 288; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MSSSNWIDAFTEELLAIDIAEASYSRSSSSSSAAPTVOATTSVHGHEPDNPQIPN 60
|||||

```
Db 1 MSSSNMIDAFTEEBELLADAIEASYNFSSSSSSSAFTVOATTSVHGHEEDPNQIPN 60
QY 61 NNRQLPRSTSTSYKRRPLSRCRARNFPAMRFGGRILYKSKATEVDKRAMOLIKYLDT 120
Db 61 NNRQLPRSTSTSYKRRPLSRCRARNFPAMRFGGRILYKSKATEVDKRAMOLIKYLDT 120
QY 121 KRDESGIAFYGLDIEMRPSFRKGVLPKQATVOICVDSNCDVMHIFHSGIPQSLQHLIE 180
Db 121 KRDESGIAFYGLDIEMRPSFRKGVLPKQATVOICVDSNCDVMHIFHSGIPQSLQHLIE 180
QY 181 DSTLVVVGIGIDGDSVKLFHDYGVSIKDYEDLSDLANOKIGDKKMGKGLASTETLVCKEL 240
Db 181 DSTLVVVGIGIDGDSVKLFHDYGVSIKDYEDLSDLANOKIGDKKMGKGLASTETLVCKEL 240
QY 241 LKPNRIRLGNWEFYPPLSKOOLQYAATDAVASMHLKYVLKDLDPVAVSGS 288
Db 241 LKPNRIRLGNWEFYPPLSKOOLQYAATDAVASMHLKYVLKDLDPVAVSGS 288
```

```
RESULT 2
US-09-896-186b-24
; Sequence 24, Application US/09896186B
; GENERAL INFORMATION:
; APPLICANT: Joshua Z. Levin
; APPLICANT: Ken Phillips
; APPLICANT: Greg Budziszewski
; APPLICANT: Fred Melns
; APPLICANT: Zhenya Glazov
; TITLE OF INVENTION: Methods of Controlling Gene Expression
; FILE REFERENCE: PB/5-31481A
; CURRENT APPLICATION NUMBER: US/09/896, 186B
; CURRENT FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 288
; TYPE: PRN
; ORGANISM: Arabidopsis thaliana
US-09-896-186b-24
```

```
Query Match 100.0%; Score 1491; DB 22; Length 288;
Best Local Similarity 100.0%; Pred. No. 6, 7e-154;
Matches 288; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSSNMIDAFTEEBELLADAIEASYNFSSSSSSSAFTVOATTSVHGHEEDPNQIPN 60
Db 1 MSSSNMIDAFTEEBELLADAIEASYNFSSSSSSSAFTVOATTSVHGHEEDPNQIPN 60
QY 61 NNRQLPRSTSTSYKRRPLSRCRARNFPAMRFGGRILYKSKATEVDKRAMOLIKYLDT 120
Db 61 NNRQLPRSTSTSYKRRPLSRCRARNFPAMRFGGRILYKSKATEVDKRAMOLIKYLDT 120
QY 121 KRDESGIAFYGLDIEMRPSFRKGVLPKQATVOICVDSNCDVMHIFHSGIPQSLQHLIE 180
Db 121 KRDESGIAFYGLDIEMRPSFRKGVLPKQATVOICVDSNCDVMHIFHSGIPQSLQHLIE 180
QY 181 DSTLVVVGIGIDGDSVKLFHDYGVSIKDYEDLSDLANOKIGDKKMGKGLASTETLVCKEL 240
Db 181 DSTLVVVGIGIDGDSVKLFHDYGVSIKDYEDLSDLANOKIGDKKMGKGLASTETLVCKEL 240
QY 241 LKPNRIRLGNWEFYPPLSKOOLQYAATDAVASMHLKYVLKDLDPVAVSGS 288
Db 241 LKPNRIRLGNWEFYPPLSKOOLQYAATDAVASMHLKYVLKDLDPVAVSGS 288
```

```
RESULT 3
US-09-513-996A-67236
; Sequence 67236, Application US/09513996A
; GENERAL INFORMATION:
; APPLICANT: N. Alexandrov et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THEREBY
; FILE REFERENCE: 2750-709P
```

```
; CURRENT APPLICATION NUMBER: US/09/513, 996A
; CURRENT FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 81028
; SEQ ID NO 67236
; LENGTH: 298
; TYPE: PRN
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 1..298
; OTHER INFORMATION: any n or Xaa = unknown
; OTHER INFORMATION: Location 1..298 / Ceres Seq. ID 2177598
US-09-513-996A-67236
```

```
Query Match 100.0%; Score 1491; DB 19; Length 298;
Best Local Similarity 100.0%; Pred. No. 7, 1e-154;
Matches 288; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSSNMIDAFTEEBELLADAIEASYNFSSSSSSSAFTVOATTSVHGHEEDPNQIPN 60
Db 11 MSSSNMIDAFTEEBELLADAIEASYNFSSSSSSSAFTVOATTSVHGHEEDPNQIPN 70
QY 61 NNRQLPRSTSTSYKRRPLSRCRARNFPAMRFGGRILYKSKATEVDKRAMOLIKYLDT 120
Db 71 NNRQLPRSTSTSYKRRPLSRCRARNFPAMRFGGRILYKSKATEVDKRAMOLIKYLDT 130
QY 121 KRDESGIAFYGLDIEMRPSFRKGVLPKQATVOICVDSNCDVMHIFHSGIPQSLQHLIE 180
Db 131 KRDESGIAFYGLDIEMRPSFRKGVLPKQATVOICVDSNCDVMHIFHSGIPQSLQHLIE 190
QY 181 DSTLVVVGIGIDGDSVKLFHDYGVSIKDYEDLSDLANOKIGDKKMGKGLASTETLVCKEL 240
Db 191 DSTLVVVGIGIDGDSVKLFHDYGVSIKDYEDLSDLANOKIGDKKMGKGLASTETLVCKEL 250
QY 241 LKPNRIRLGNWEFYPPLSKOOLQYAATDAVASMHLKYVLKDLDPVAVSGS 288
Db 251 LKPNRIRLGNWEFYPPLSKOOLQYAATDAVASMHLKYVLKDLDPVAVSGS 298
```

```
RESULT 4
US-09-708-427-27138
; Sequence 27138, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. Alexandrov et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708, 427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 27138
; LENGTH: 313
; TYPE: PRN
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..313
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: 1..313
; OTHER INFORMATION: Ceres Seq. ID 1820253
US-09-708-427-27138
```

```
Query Match 96.5%; Score 1439.5; DB 21; Length 313;
Best Local Similarity 95.3%; Pred. No. 3, 4e-148;
Matches 281; Conservative 1; Mismatches 4; Indels 9; Gaps 1;

QY 1 MSSSNMIDAFTEEBELLADAIEASYNFSSSSSSSAFTVOATTSVHGHEEDPNQIPN 60
Db 1 MSSSNMIDAFTEEBELLADAIEASYNFSSSSSSSAFTVOATTSVHGHEEDPNQIPN 60
```



```

; LOCATION: 1..197
; OTHER INFORMATION: any n or Xaa - unknown
; FEATURE:
; OTHER INFORMATION: Location 1..197 / Ceres Seq. ID 1572653
US-09-513-996A-28857

Query Match      66.8%; Score 996; DB 19; Length 197;
Best Local Similarity 97.0%; Pred. No. 6.9e-100;
Matches 191; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Db
QY 92 MREGRIILYKTEVDKRAMOLIKVLDTRKRDSEGIATVGLDIEMRSPFRKGYLPKGVAT 151
   1 MREGRIILYKTEVDKRAMOLIKVLDTRKRDSEGIATVGLDIEMRSPFRKGYLPKGVAT 60
QY 152 VOICVDSNCDVWHIFHSGIPQSLQHLIEDSTLVKVGIGIDGSVKLFHDYGVSIKDED 211
   61 VOICVDSNCDVWHIFHSGIPQSLQHLIEDSTLVKVGIGIDGSVKLFHDYGVSIKDED 120
Db 212 LSDLANOKTIGDCKKGLASTETLVCKELLPRIIRLGNWMEFYPLSKOOLQYAATDAYAS 271
   121 LSDLANOKTIGDCKKGLASTETLVCKELLPRIIRLGNWMEFYPLSKOOLQYAATDAYAS 180
QY 272 WHLYKVLKDLPPAVSGS 288
   181 WHLYKVLKDLPPAVSGS 197

RESULT 9
US-09-513-996A-28857
; Sequence 28857, Application US/09513996A
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-709P
; CURRENT APPLICATION NUMBER: US/09/513, 996A
; CURRENT FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 81028
; SEQ ID NO 28857
; LENGTH: 197
; TYPE: PRP
; ORGANISM: Arabidopsis thaliana
; NAME/KEY: UNSURE
; LOCATION: 1..197
; OTHER INFORMATION: any n or Xaa - unknown
; FEATURE:
; OTHER INFORMATION: Location 1..197 / Ceres Seq. ID 2177600
US-09-513-996A-67238

Query Match      68.9%; Score 1028; DB 19; Length 197;
Best Local Similarity 100.0%; Pred. No. 2.1e-103;
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db
QY 92 MREGRIILYKTEVDKRAMOLIKVLDTRKRDSEGIATVGLDIEMRSPFRKGYLPKGVAT 151
   1 MREGRIILYKTEVDKRAMOLIKVLDTRKRDSEGIATVGLDIEMRSPFRKGYLPKGVAT 60
QY 152 VOICVDSNCDVWHIFHSGIPQSLQHLIEDSTLVKVGIGIDGSVKLFHDYGVSIKDED 211
   61 VOICVDSNCDVWHIFHSGIPQSLQHLIEDSTLVKVGIGIDGSVKLFHDYGVSIKDED 120
Db 212 LSDLANOKTIGDCKKGLASTETLVCKELLPRIIRLGNWMEFYPLSKOOLQYAATDAYAS 271
   121 LSDLANOKTIGDCKKGLASTETLVCKELLPRIIRLGNWMEFYPLSKOOLQYAATDAYAS 180
QY 272 WHLYKVLKDLPPAVSGS 288
   181 WHLYKVLKDLPPAVSGS 197
```

```

; LOCATION: 1..197
; OTHER INFORMATION: any n or Xaa - unknown
; FEATURE:
; OTHER INFORMATION: Location 1..197 / Ceres Seq. ID 1572653
US-09-513-996A-28857

Query Match      66.8%; Score 996; DB 19; Length 197;
Best Local Similarity 97.0%; Pred. No. 6.9e-100;
Matches 191; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Db
QY 92 MREGRIILYKTEVDKRAMOLIKVLDTRKRDSEGIATVGLDIEMRSPFRKGYLPKGVAT 151
   1 MREGRIILYKTEVDKRAMOLIKVLDTRKRDSEGIATVGLDIEMRSPFRKGYLPKGVAT 60
QY 152 VOICVDSNCDVWHIFHSGIPQSLQHLIEDSTLVKVGIGIDGSVKLFHDYGVSIKDED 211
   61 VOICVDSNCDVWHIFHSGIPQSLQHLIEDSTLVKVGIGIDGSVKLFHDYGVSIKDED 120
Db 212 LSDLANOKTIGDCKKGLASTETLVCKELLPRIIRLGNWMEFYPLSKOOLQYAATDAYAS 271
   121 LSDLANOKTIGDCKKGLASTETLVCKELLPRIIRLGNWMEFYPLSKOOLQYAATDAYAS 180
QY 272 WHLYKVLKDLPPAVSGS 288
   181 WHLYKVLKDLPPAVSGS 197

RESULT 10
US-09-708-427-27139
; Sequence 27139, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708, 427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 27139
; LENGTH: 222
; TYPE: PRP
; ORGANISM: Arabidopsis thaliana
; NAME/KEY: misc.feature
; LOCATION: 1..222
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: 1..222
; OTHER INFORMATION: Ceres Seq. ID 1820254
US-09-708-427-27139

Query Match      65.5%; Score 976.5; DB 21; Length 222;
Best Local Similarity 93.1%; Pred. No. 1.2e-97;
Matches 190; Conservative 1; Mismatches 4; Indels 9; Gaps 1;

Db
QY 92 MREGRIILYKTEVDKRAMOLIKVLDTRKRDSEGIATVGLDIEMRSPFRKGYLPKGVAT 151
   1 MREGRIILYKTEVDKRAMOLIKVLDTRKRDSEGIATVGLDIEMRSPFRKGYLPKGVAT 60
QY 152 VOICVDSNCDVWHIFHSGIPQSLQHLIEDSTLVKVGIGIDGSVKLFHDYGVSIKDED 211
   61 VOICVDSNCDVWHIFHSGIPQSLQHLIEDSTLVKVGIGIDGSVKLFHDYGVSIKDED 120
Db 212 LSDLANOKTIGDCKKGLASTETLVCKELLPRIIRLGNWMEFYPLSKOOLQYAATDAYAS 271
   121 LSDLANOKTIGDCKKGLASTETLVCKELLPRIIRLGNWMEFYPLSKOOLQYAATDAYAS 180
QY 272 WHLYKVLKDLPPAVSGS 286
   181 WHLYKVLKDLPPAVSGS 204
```

Query Match	58.8%	Score	876.5	DB	21	Length	202
Best Local Similarity	92.4%	Pred. No.	9	Le	87		
Matches	170	Conservative	1	Mismatches	4	Indels	9
Gaps							1
QY	112	MQLIKVLDTKDESGIAFVGLDIEMRPSFRKGLVPEKKAATVQICVDSNYCDVMHIFHSGI	171				
Db	1	MQLIKVLDTKDESGIAFVGLDIEMRPSFRKGLVPEKKAATVQICVDSNYCDVMHIFHSGI	60				

```

, RESULT 14
, US-09-906-226-16
, Sequence 16, Application US/09906226
, GENERAL INFORMATION:
, APPLICANT: Butler, Karla
, APPLICANT: Cahoon, Rebecca E.
, APPLICANT: Rafalski, Antoni
, APPLICANT: Sakai, Hajime
, TITLE OF INVENTION: Plant Rasad-Like Genes
, FILE REFERENCE: B1467 US NA
, CURRENT APPLICATION NUMBER: US/09/906.226
, CURRENT FILING DATE: 2001-07-16
, PRIOR APPLICATION NUMBER: 60/218993
, PRIOR FILING DATE: July 17, 2000
, NUMBER OF SEQ ID NOS: 57
, SOFTWARE: Microsoft Office 97
, SEQ ID NO 16

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 10:23:40 : Search time 41 Seconds
(without alignments)
1513.077 Million cell updates/sec

Title: US-09-896-186B-24

Perfect score: 1491

Sequence: 1 MSSSNWIDAFTEBELLAID.....YASWHLKVKLKDLPDAVSGS 288

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1023978 seqs, 215403328 residues

Total number of hits satisfying chosen parameters: 1023978

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/2/paa/PCF_NEW_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	666.5	44.7	268	6	US-10-424-599-227379 Sequence 227379, A
2	352	23.6	994	5	US-09-724-676-89006 Sequence 89006, A
3	352	23.6	994	5	US-09-724-676A-89006 Sequence 89006, A
4	352	23.6	1409	5	US-09-949-001-22 Sequence 22, Appl
5	352	23.6	1432	5	US-09-949-001-16 Sequence 16, Appl
6	352	22.1	1432	6	US-10-374-077-71 Sequence 71, Appl
7	330	22.1	1401	6	US-10-374-077-206 Sequence 206, Appl
8	253.5	17.0	238	6	US-10-424-599-238476 Sequence 238476, A
9	249.5	16.7	210	6	US-10-425-114-37706 Sequence 37706, A
10	245	16.4	236	6	US-10-425-114-71107 Sequence 71107, A
11	212	14.2	208	6	US-10-424-599-191095 Sequence 191095, A
12	188	12.6	205	6	US-10-424-599-229443 Sequence 229443, A
13	169.5	11.4	223	6	US-10-431-652-6289 Sequence 6289, Ap
14	163.5	11.0	198	6	US-10-425-114-71315 Sequence 71315, A
15	151	10.1	127	6	US-10-276-774-1634 Sequence 1634, Ap
16	142	9.5	599	6	US-10-094-749-1828 Sequence 1828, Ap
17	119.5	8.0	202	6	US-10-424-599-254772 Sequence 254772, A
18	117	7.8	511	6	US-10-425-114-43332 Sequence 43332, A
19	99	6.6	126	6	US-10-424-599-220783 Sequence 220783, A
20	94.5	6.3	159	6	US-10-424-599-274357 Sequence 274357, A
21	93	6.2	488	6	US-10-369-493-18241 Sequence 18241, A
22	93	6.2	1348	5	US-09-949-002-517 Sequence 517, Appl
23	93	6.2	2207	5	US-09-724-676-60068 Sequence 60068, A
24	93	6.2	2207	5	US-09-724-676A-60068 Sequence 60068, A
25	93	6.2	2224	1	PCT-US02-19017-31 Sequence 31, Appl
26	93	6.2	2224	5	US-09-949-002-292 Sequence 292, Appl

27	93	6.2	2224	7	US-60-452-680-22968 Sequence 22968, A
28	93	6.2	2224	7	US-60-453-135-14275 Sequence 14275, A
29	93	6.2	2224	7	US-60-453-050-14275 Sequence 14275, A
30	93	6.2	2224	7	US-60-455-444-7663 Sequence 7663, Ap
31	93	6.2	2224	7	US-60-465-241-7663 Sequence 7663, Ap
32	93	6.2	2224	7	US-60-466-412-14275 Sequence 14275, A
33	93	6.2	3079	6	US-10-369-493-2024 Sequence 2024, Ap
34	92.5	6.2	481	6	US-10-431-652-4475 Sequence 4475, Ap
35	91.5	6.1	1836	5	US-09-949-016-7432 Sequence 7432, Ap
36	89.5	6.0	357	6	US-10-417-886-8398 Sequence 8398, Ap
37	89.5	6.0	930	6	US-10-282-122A-58327 Sequence 58327, A
38	89.5	6.0	1324	6	PCT-US02-32637-53 Sequence 53, Appl
39	89.5	6.0	1324	6	US-10-270-839-53 Sequence 53, Appl
40	89.5	6.0	1324	7	US-60-427-165-32 Sequence 32, Appl
41	89	6.0	489	1	PCT-US02-36123-2614 Sequence 2614, Ap
42	89	6.0	560	1	PCT-US03-06415-2 Sequence 2, Appl1
43	89	6.0	892	6	US-10-282-122A-70481 Sequence 70481, A
44	89	6.0	930	6	US-10-092-411A-5314 Sequence 5314, Ap
45	88.5	5.9	480	6	US-10-424-599-262034 Sequence 262034, A

ALIGNMENTS

```
RESULT 1
US-10-424-599-227379
; Sequence 227379, Application US/10424599
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21153223B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285084
; SEQ ID NO 227379
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_47352C.1.pep
US-10-424-599-227379

Query Match          44.7% Score 666.5; DB 6; Length 268;
Best Local Similarity 51.1% Pred. No. 9.4e-51;
Matches 134; Conservative 46; Mismatches 61; Indels 21; Gaps 5;

QY      29 SRSSSSSSAAPVQATTSVHGHEEDPNOIRNIRLOPNSI-----TSSTSYKRPPLSR 83
      6 SEASLSNNKKRP-----FNDHTHTP-----RRRLPKSLIALQHPNASSFSRPP-RP 50
      84 CRAR-NFPMRFGGRTKATEYDKRAMOLIKYLDTRKRDSEGIAPFGIIEWRPSRK 142
      51 CDSRRTTLYVKFSGSISRTKPDAYEKATLLOLQLEKTTDMQOTALGFDEMKVYRK 110
      143 GVLPGKAVTVOICVDSNYCDMHLFHSQIPSOHLHEDSTLVYVGIGIDSDSVKLFHDY 202
      111 GVPPGKAVVMQICGFRCHVHLHLSITPNTLOLLEDPYLVKAGNIDDDAVKVFHDY 170
      171 NISVGYVDLSEFHNQKIGCDHKWGLASLFTFKLSKQKPKNKIRLGNWEAPVLSKQLE 230
QY      203 GVSIRKVEDSLDANOKTIGSKKWLGLASLFTFKLSKQKPKNKIRLGNWEAPVLSKQLE 262
      263 YAATDAVSMHLKVKIKDLPDA 284
      231 YAATDAVSMCLYQIKDLPDA 252
      Db
      RESULT 2
      US-09-724-676-89006
```

```
; Sequence 89006, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 89006
; LENGTH: 994
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-89006
```

```
Query Match      23.6%; Score 352; DB 5; Length 994;
Best Local Similarity 37.7%; Pred. No. 3.5e-22;
Matches 75; Conservative 41; Mismatches 75; Indels 8; Gaps 3;
```

```
QY 88 NPPAMRFGRIILYKATATEVDKRAMOLIKVLDTRKDESGIAFYGLDIEMRPSFRKGVLP 147
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 43 DLPLEFTGSIYSDASCSFLSE-----DISMSISDGDVYGFPMEMPPLYNRGL-G 95

QY 148 KATVQICVDNSNCVDMHIFHSGI--POSLOHLIEDSTLVKVGIGIDGDSVKLFHDYGVSI 206
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 96 KVALIQLCVSESKCYLFHVSMSSVFPQGLKMLENNAVKAGVIGIGIDGDKLLRDPDIKL 155

QY 207 KDVEDSLDANOKIGGDKKWLASLTETLVCKELLPNRIIRLGNMFEFPLSKOOLQYAT 266
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 156 KNFVELTDVANKKLTCTETWSLNSLVKHLGKQLLDKSIKCSNWSKFPLETEDOKLYAAT 215

QY 267 DAYASWHLKYVLKDLDPAY 285
   |||| : : : : : : : : : :
Db 216 DAYAGFIIRNLEIILDDTV 234
```

```
RESULT 3
US-09-724-676A-89006
; Sequence 89006, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 89006
; LENGTH: 994
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-89006
```

```
Query Match      23.6%; Score 352; DB 5; Length 994;
Best Local Similarity 37.7%; Pred. No. 3.5e-22;
Matches 75; Conservative 41; Mismatches 75; Indels 8; Gaps 3;
```

```
QY 88 NPPAMRFGRIILYKATATEVDKRAMOLIKVLDTRKDESGIAFYGLDIEMRPSFRKGVLP 147
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 43 DLPLEFTGSIYSDASCSFLSE-----DISMSISDGDVYGFPMEMPPLYNRGL-G 95

QY 148 KATVQICVDNSNCVDMHIFHSGI--POSLOHLIEDSTLVKVGIGIDGDSVKLFHDYGVSI 206
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 96 KVALIQLCVSESKCYLFHVSMSSVFPQGLKMLENNAVKAGVIGIGIDGDKLLRDPDIKL 155

QY 207 KDVEDSLDANOKIGGDKKWLASLTETLVCKELLPNRIIRLGNMFEFPLSKOOLQYAT 266
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 156 KNFVELTDVANKKLTCTETWSLNSLVKHLGKQLLDKSIKCSNWSKFPLETEDOKLYAAT 215

QY 267 DAYASWHLKYVLKDLDPAY 285
   |||| : : : : : : : : : :
Db 216 DAYAGFIIRNLEIILDDTV 234
```

```
RESULT 4
US-09-949-001-22
; Sequence 22, Application US/09949001
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL000789
; CURRENT APPLICATION NUMBER: US/09/949,001
; PRIOR FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/231,323
; NUMBER OF SEQ ID NOS: 848
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 1409
; TYPE: PRT
; ORGANISM: Human
US-09-949-001-22
```

```
Query Match      23.6%; Score 352; DB 5; Length 1409;
Best Local Similarity 37.7%; Pred. No. 5.5e-22;
Matches 75; Conservative 41; Mismatches 75; Indels 8; Gaps 3;
```

```
QY 88 NPPAMRFGRIILYKATATEVDKRAMOLIKVLDTRKDESGIAFYGLDIEMRPSFRKGVLP 147
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 47 DLPLEFTGSIYSDASCSFLSE-----DISMSISDGDVYGFPMEMPPLYNRGL-G 99

QY 148 KATVQICVDNSNCVDMHIFHSGI--POSLOHLIEDSTLVKVGIGIDGDSVKLFHDYGVSI 206
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 100 KVALIQLCVSESKCYLFHVSMSSVFPQGLKMLENNAVKAGVIGIGIDGDKLLRDPDIKL 159

QY 207 KDVEDSLDANOKIGGDKKWLASLTETLVCKELLPNRIIRLGNMFEFPLSKOOLQYAT 266
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 160 KNFVELTDVANKKLTCTETWSLNSLVKHLGKQLLDKSIKCSNWSKFPLETEDOKLYAAT 219

QY 267 DAYASWHLKYVLKDLDPAY 285
   |||| : : : : : : : : : :
Db 220 DAYAGFIIRNLEIILDDTV 238
```

```
RESULT 5
US-09-949-001-16
; Sequence 16, Application US/09949001
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL000789
; CURRENT APPLICATION NUMBER: US/09/949,001
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/231,323
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 848
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 1432
; TYPE: PRT
; ORGANISM: Human
US-09-949-001-16
```

```
Query Match      23.6%; Score 352; DB 5; Length 1432;
Best Local Similarity 37.7%; Pred. No. 5.7e-22;
Matches 75; Conservative 41; Mismatches 75; Indels 8; Gaps 3;
```

```
QY 88 NPPAMRFGRIILYKATATEVDKRAMOLIKVLDTRKDESGIAFYGLDIEMRPSFRKGVLP 147
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 43 DLPLEFTGSIYSDASCSFLSE-----DISMSISDGDVYGFPMEMPPLYNRGL-G 95

QY 148 KATVQICVDNSNCVDMHIFHSGI--POSLOHLIEDSTLVKVGIGIDGDSVKLFHDYGVSI 206
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 96 KVALIQLCVSESKCYLFHVSMSSVFPQGLKMLENNAVKAGVIGIGIDGDKLLRDPDIKL 155
```



```

; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 238476
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_57369C.1.pep
US-10-424-599-238476
```

```

Query Match      17.0%; Score 253.5; DB 6; Length 238;
Best Local Similarity 39.9%; Pred. No. 2.6e-14;
Matches 59; Conservative 28; Mismatches 56; Indels 5; Gaps 4;
```

```

QY 130 VGLDIEMRSPFRKGVLPKAVATQICVDSNYCDVMHIFHS-GIPQSLQHLIEDSTLVKVG 188
    |||||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 85 VGLDIEMRPTQRN-QMNPVATIQLCV-AERCLVFQILHSPSPISPLVSLPADPNITFVG 142
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 189 IGIDGSKLFIHYGVSIKDYEDLSLANOKIG--DKKGLASLFTETLVCKELKPNRI 246
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 143 VGIOEVEKLEEDYNINAVNRDLRSFAERLGDLEIKRAGLKSGLRVLGLEVAKPKRY 202
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 247 RLGNWEFYPLSKOOLQOYAATDAYASWHL 274
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 203 TRSRMDPMLTAAQOVYAAVDAFLSYEI 230
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```

RESULT 9

```

US-10-425-114-37706
; Sequence 37706, Application US/10425114
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 37706
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700346357_FLI.pep
US-10-425-114-37706
```

```

Query Match      16.7%; Score 249.5; DB 6; Length 210;
Best Local Similarity 28.7%; Pred. No. 5.5e-14;
Matches 58; Conservative 45; Mismatches 90; Indels 9; Gaps 5;
```

```

QY 80 PLSCARARFPMPFGRRIYKTAIEVDKRAMOLIKVLTQRDESGIAFVGLDIEMRPS 139
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 10 PMATATVCN--VRFEGNVITTVTASGAAVESMDELISVHRRRLIKLVGGLDVEWRPS 66
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 140 FRKGVLPKAVATQICVDSNYCDVMHIFHS-GIPQSLQHLIEDSTLVKVGIGIDGSKVL 198
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 67 FSNAY--SKTATVQLCV-GRCLIFQLHADVVPNTLDEFLSDPDYTFVGVAADVERL 123
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 199 FHDYGVSIKDYEDLSLANOKIGD--KKWGLASLFTETLVCKELKPNRIKGNWEFYPL 256
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 124 ENYIDLEVANAEDLAEKAKEMGRPLRNAGLOGIARAVDAHVEKRPQWRTGPMWASSL 183
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```

```

QY 257 SKOOLQYAATDAYASWHLKVL 278
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 184 SDEQIEYATIDAFVSFEVGRML 205
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```

RESULT 10

```

US-10-425-114-71107
; Sequence 71107, Application US/10425114
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 71107
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4762-022-E12_FLI.pep
US-10-425-114-71107
```

```

Query Match      16.4%; Score 245; DB 6; Length 236;
Best Local Similarity 28.9%; Pred. No. 1.6e-13;
Matches 55; Conservative 43; Mismatches 86; Indels 6; Gaps 4;
```

```

QY 92 MRFGRITYSKTATEVDKRAMOLIKVLTQRDESGIAFVGLDIEMRSPFRKGVLPKAVAT 151
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 45 VREGANVITTVTASGAAVESMDELISVHRRRLIKLVGLDVEWRPSFSRAY--SKTAT 102
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 152 VOICVDSNYCDVMHIFHS-GIPQSLQHLIEDSTLVKVGIGIDGSKLFIHYGVSIKDY 210
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 103 LQLCV-GRCLIFQLHADVVPNTLDEFLSDPDYTFVGVAADVERLENDYDLEVANA 161
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 211 DLSLANOKIGD--KKWGLASLFTETLVCKELKPNRIKGNWEFYPLSKOOLQYAATDA 268
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 162 DLAEKAKEMGRPLRNAGLOGIARAVDAHVEKRPQWRTGPMWASSLSDEQIEYATIDA 221
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 269 YASWHLKVL 278
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 222 FVSFEVGRML 231
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```

RESULT 11

```

US-10-424-599-191095
; Sequence 191095, Application US/10424599
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 191095
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_14578C.1.pep
US-10-424-599-191095
```

```

Query Match      14.2%; Score 212; DB 6; Length 208;
```

Best Local Similarity 31.7%; Pred. No. 1.1e-10;
Matches 58; Conservative 31; Mismatches 60; Indels 34; Gaps 6;

QY 120 TKRDSGTA-----FVGGLDIEMRSPFRKGYLPGRKATVQICVDSNYCYD 163
Db 32 TTDSGTVDKMIOVVSSTYAGKORIVGIDTETWTKPKRM--KVAIIQLCE-NKCL 87
QY 164 MHIFH-SGIPQSLQHLIEDSTLVKVGIGIDGSKLPHDYGVSIKDVEDLSDLANOKIG 222
Db 88 IOLFMDNIPOSLRSLMDSNFEVGVGYINDLRMLKNYGLECKNGIDVSLAKE----- 143
QY 223 DKW-----GLASITFVLCKELKPNRIRLGNMEFYPLSKQOLQYATDAYASWHL 275
Db 144 --KMPHRISSGALKYLAKEVLEMEKSKAVCTSEWQSKELQTOIEVACIDAYASFKIG 201
QY 276 KVL 278
Db 202 KMI 204

RESULT 12

US-10-424-599-229443
; Sequence 229443, Application US/10424599
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424, 599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 229443
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_49211C.1.pep
US-10-424-599-229443

Query Match
Best Local Similarity 60.4%; Pred. No. 1.5e-08;
Matches 32; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

QY 152 VOICVDSNYCDVMHIFHSGIPQSLQHLIEDSTLVKVGIGIDGSKLPHDYG 204
Db 1 MOICGDTTRCHVTLHLSGIPQNTQLLEDPVTLKVGAGIDGDAVVFDDYNI 53

RESULT 13

US-10-431-652-6289
; Sequence 6289, Application US/10431652
; GENERAL INFORMATION:
; APPLICANT: Breton, Gary L.
; APPLICANT: Bush, David
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNIT FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PAT03-08
; CURRENT APPLICATION NUMBER: US/10/431, 652
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: US 09/328, 352
; PRIOR FILING DATE: 1999-06-04
; PRIOR APPLICATION NUMBER: US 60/088, 701
; PRIOR FILING DATE: 1998-06-09
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6289
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-10-431-652-6289

Query Match
Best Local Similarity 11.4%; Score 169.5; DB 6; Length 223;
Matches 45; Conservative 35; Mismatches 72; Indels 11; Gaps 5;

QY 129 FVGGLDIEMRSPFRKGYLPGRKATVQICVDSNYCDVMHIFHSGIP--QSLQHLIEDSTLVK 186
Db 58 FLGFDESKPTFQVGVSTGPHLIQLATEHK---AYLFVNLSLTKFLQPILSNPKQIK 113
QY 187 VGIQIDGDSVKLPHDYGVSIKDVEDLSDLANOKIGDKKMGKGLASLETETVCKELKPNRI 246
Db 114 VGFGLKNDK-HIFHKGIELESCVDLAK-GFSHFGRQOMGVOKAVALLGOYLYANSKV 171
QY 247 RLGNMEFYPLSKQOLQYATDAYASWHLKVD--LPDAYS 286
Db 172 GTSNMARKPLTSQISYAAADATALLVFLLELKKRVLPRIHS 214

RESULT 14

US-10-425-114-71315
; Sequence 71315, Application US/10425114
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425, 114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 71315
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73153A09_FLI.pep
US-10-425-114-71315

Query Match
Best Local Similarity 28.5%; Pred. No. 2.1e-06;
Matches 47; Conservative 33; Mismatches 74; Indels 11; Gaps 5;

QY 118 LDTKRDSGTAFLVGLDIEMRSPFRKGYLPGRKATVQICVD--SNYCDVMHIFH---SGIP 172
Db 1 LNATSDIEACKIIGMCEMPNFEKWTSSKSVIIQIASDKIAFIPLDLIKYEDDPKALD 60
QY 173 QSLQHLIEDSTLVKVGIGIDGSKLPHDYG--VSIKDVEDLSDLANOKIGDKKMGKGLAS 230
Db 61 SCILRRVWSSKILKLYDIQCDLHQLTRSYGLECFOSYEMVLDM--QKLFKVTGGLSG 118
QY 231 LFTLVCKELKPNRIRLGNMEFYPLSKQOLQYATDAYASWHL 275
Db 119 LSKETILGAGLNKTR--RNSWNEORPLTONKETAALDAVYVHIF 161

RESULT 15

US-10-276-774-1634
; Sequence 1634, Application US/10276774
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y, Tom et al
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276, 774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560, 875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496, 914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700

Sat Jun 14 11:07:48 2003

us-09-896-186b-24.rapn

Page 6

```

; SOFTWARE: Custom
; SEQ ID NO 1634
; LENGTH: 127
; TYPE: PRT
; ORGANISM: Homo
US-10-276-774-1634

```

ORGANISM: Homo sapiens
US-10-276-774-1634

Query Match	10.1%;	Score 151;	DB 6;	Length 127;
Best Local Similarity	35.0%;	Pred. No. 1.5e-05;		
Matches	41;	Conservative	20;	Mismatches 48;
			Indels	8;
			Gaps	2;

```

QY      171  IPSTLHLLIEDSTLKVGIGIDGDSVKLFHHYGVSFKDVEDLSDLA-----NQKIGCDMK   225
          ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      3    LPRLLDIADLTILKVGCSSEDAKSLLDQRYGLVVRGCCLDRYLAMRQRNNLLCNG---  59

```

QY 226 WGLASLTETLVCKELLKPNRRRLGNWFYPLSKOOLQYAATDAYASWHLYKVLKDLR 282
| | | | | : : | | : : | | | | : : |
Db 60 LSLKSAETVLPNFPIDKSLLRCSNMDAETLTEDQVIYAARDQAISVALFLHLGLYP 116

Search completed: June 6, 2003, 10:31:03
Job time is: 43 secs